DOCUMENT RESUME

ED 365 677 TM 019 283

AUTHOR	Thomas, Brenda; Badger, Elizabeth
TITLE	On Their Own: Students' Responses to Open-Ended Questions in Math, Reading, Science, Social Studies. Results of the 1990 Assessment: Grade Four, Grade Eight, Grade Twelve.
INSTITUTION	Massachusetts State Dept. of Education, Boston. Massachusetts Educational Assessment Program.
PUB DATE	Apr 91
NOTE	169p.
PUB TYPE	Statistical Data (110) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Educational Assessment; Elementary School Students; Elementary Secondary Education; Grade 4; Grade 8; Intermediate Grades; Junior High School Students; *Mathematics Tests; *Reading Tests; *Responses; *Science Tests; *Social Studies; *State Programs; Statistical Data; Testing Programs; Test Results
IDENTIFIERS	Massachusetts Educational Assessment Program; *Open Ended Questions

ABSTRACT

In the spring of 1990, over 9,000 4th-, 8th- and 12th-grade students (only 6 percent of all students in Massachusetts) were assessed using open-ended mathematical, scientific, social studies, and reading concepts. Beginning with the Massachusetts Educational Assessment Program for 1992, open-ended questions will be administered to all students and will contribute to school and district scores. This series of reports describes the results of these assessments to communicate levels of student achievement throughout the state, familiarize teachers and administrators with the types of questions that will be included on the next assessment; and improve assessments taking place within classrooms by providing models that teachers can adapt to their own evaluations of students' knowledge, understanding, and abilities. Fifteen handouts for grade 4, 17 handouts for grade 8, and 15 handouts for grade 12, are provided. Students' responses to open-ended questions in social studies, mathematics, science, and reading are highlighted. For selected test items in each subject area, the grade level, major concepts/abilities tested, correct and incorrect answers, and comments are included. (RLC)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED 365 677

ON THEIR OWN

Students' Responses to Open-Ended Questions In Math, Reading, Science, Social Studies

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. RAYNOLDS, JR.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Results of the 1990 Assessment

Grade Four

Massachusetts Department of Education

Massachusetts Educational Assessment Program

TM019283

BEST COPY AVAILABLE

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Social Studies Background

Currently, social studies is experiencing a "back to basics" movement stressing history, civics, and geography. This trend portends new curriculum examination and revitalization. Recent research has indicated that in elementary and middle grades, teaches make false assumptions about what their students know. Textbooks contribute to this problem by making the same assumption. In a recent article by Margaret McKeown and Isabel Beck,[1] researchers considered the background knowledge that middle school students bring to the study of history, the American Revolution in particular. They say that

> "elementary social studies textbooks seem to assume knowledge of more specific concepts than students have under control...students do seem to have a general landscape involving the country being settled, the seeking of freedom and some notions of a process...Yet this seems exactly where the texts focus their material."

This miscalculation of student knowledge calls for more complete and in-depth assessments of what significant knowledge is lacking. The longer that misconceptions continue uncorrected, the more difficult it becomes to correct them. Students will integrate knowledge to accommodate misinformation.

According to McKeown and Beck, one of the greatest challenges facing educators is finding and correcting

the confusions that young students bring to class. They suggest using a framework of the topic to organize previous knowledge and new information into a coherent whole. By presenting the relevant information in a framework, students would be better able to see the totality of the topic and evaluate whether their previous knowledge will fit in. They might be able to recognize irrelevant or inaccurate information using the framework as a sounding board.

Retention of knowledge is a problem at the secondary level. Social studies must build on the knowledge structures of the past. All this calls for a unified curriculum that considers how and when to teach materials. Walter Parker in *Renewing the Social Studies Curriculum*[2] suggests that themes in social studies can begin at the elementary grades and be reinforced across the years; that is, "content at any grade level should be presented in ways that provide, insofar as possible, a comprehensive view of a complex whole." This suggests a thematic approach without sacrificing chronology or coverage, which would satisfy those who call for social studies to have breadth over depth.

Parker also discusses attributes of a successful social studies classroom.

- Lessons have substantive coherence and continuity. Disparate facts are pulled together.
- Students are given time to think—to prepare responses to questions.
- The teacher asks challenging questions and structures challenging tasks.
- Students are able to offer explanations and reasons for their conclusions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, the open-ended questions covered process skills such as map reading and interpreting graphic representations. Another question focussed on a sense of history. Other questions involved conservation and economics.

Fourth grade teachers were given a small selection of these questions and asked to predict how prepared their students were to answer them successfully. Over half the teachers felt that their students were very well prepared to compare their own lives with those of colonial school children, while 33% of the students were able to do so. In response to how well their students were prepared to write a set of directions using a map, 31% of the teachers felt that their students were very well prepared. One-quarter of the students were able to do so.

Grade Eight

Eighth grade students were asked to interpret a political cartoon. A series of questions asked students to evaluate the quality of several arguments. Both specifics of history (students were asked to place six events on a time line) and a sense of history (students had to evaluate evidence to identify a speaker and when he lived) were covered. Students were asked to generate three courses of action and their negative impacts in response to an environmental crisis.

As in grade four, social studies teachers were given a small selection of open-ended questions and asked to evaluate how prepared their students were to answer those questions. About 15% of the teachers felt their students were very well prepared to generate three courses of action and their corresponding negative impacts; approximately 20% of the students were able to do so. When asked how well prepared their students were to explain why an argument was strong, about half the teachers felt that their students were very well prepared to do so. Eleven percent of the students were able to interpret and explain the strength of an argument.

Grade Twelve

Seniors were asked to interpret population graphs and make predictions based on those graphs. They were also asked to discuss how the policy of open immigration has affected the United States socially, politically, and economically. Another series of questions asked students to interpret two different graphs of immigration rates and discuss changes the graphs illustrated. As in grade eight, seniors were asked to interpret a political cartoon, this one about gun control.

Teachers who taught twelfth grade social studies were asked to comment on the preparedness of their students when given a sample of the questions. About one fourth (27%) of the teachers felt that their students were very well prepared to discuss the political, economic, and social effects of immigration while no more than 13% of the students were able to give complete answers in any of the categories. When given two graphs of immigration rates, one spanning 155 years, the other one year, 82% of the teachers felt their students would be able to answer why the numbers were larger in one graph than the other. Thirty-five percent of the seniors were able to answer that one graph spanned a much longer period of time than the other.

At grades eight and twelve, it appears that teachers felt their students were better prepared than they were to answer those questions that dealt with higher order thinking skills, such as evaluating arguments (grade eight) and cause and effect (effects of immigration at grade twelve).

In grades eight and twelve, the majority of teachers cited student indifference and lack of motivation as important factors in students' lack of success in social studies. Almost half the teachers at both grades cited poor preparation prior to entering the class as another important factor in students' lack of success.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

---

[1] McKeown, Margaret G. And Beck, Isabel L. (Winter 1990). "The Assessment & Characterization of Young Learners' Knowledge of a Topic in History." American Educational Research Journal. Vol. 27, 4, pp. 688-726.

[2] Parker, Walter C. (1991). ***Reviewing the Social Studies Curriculum.*** Alexandria, Virginia: Association for Supervision and Curriculum Development.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Math Background

In 1990 the National Council for Teachers of Mathematics published ***The Curriculum and Evaluation Standards***, which challenged the traditional building block model of mathematics learning. According to this model, school mathematics consists of a set of discrete skills or building blocks—computational algorithms, measurement formulae, and algebraic rules. The elementary school teacher sets the first layer of blocks, the middle school teacher the second layer, and so on until, eventually the edifice is large and strong enough for students to be said to know mathematics. Although problem solving is a goal of learning, it is not actually connected to instruction. It comes later as a natural consequence of learning. The role of the teacher is to ensure that students solidify their learning at each level in order to progress to the next until, ultimately, they would be able to "do mathematics," i.e., solve real problems.

Unfortunately, for most students, this model is not working. Students are not automatically solving problems. They are not making the connections. Furthermore, most are getting stuck at the lower levels, so they are not even allowed the chance to meet real problems.

In contrast, the Standards suggested a different model for mathematics learning—one of a network of skills and concepts that are related to one another. Students can go in and out of the network at any level because the ideas are connected. Furthermore, problems form

the context of this network, because solving problems is how we begin to understand what mathematics is about.

In their publication of the *Standards*, as well as the more recent *Professional Standards for Teaching Mathematics*, the National Council for Teachers of Mathematics stresses understanding over memory, relevance over training. It advocates an expansion of the curriculum to include new topics—such as probability and statistics, patterns and relationships, discrete mathematics—that reflect the mathematical needs of society. It gives greater emphasis to other topics—such as geometry and measurement—that have often been relegated to the back of the book. It delineates four basic abilities that are essential to all topics: reasoning, communication, problem solving, and the ability to see the connections between mathematical concepts. By including evaluation, it makes the point that assessment should be viewed as an integral part of instruction.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

The open-ended questions were designed to cover different important concepts and abilities in the fourth grade mathematics curriculum. Questions concentrated on numerical operations and the ability to extract and use information presented in graphical form. There was little stress on calculation itself. Instead, students were asked to represent multiplication and division by writing stories and drawing pictures. They were required to interpret the data that were presented in a set of graphs and were asked to organize data. A final set of questions dealt with spatial visualization and geometric shapes.

When given a small selection of the actual questions, less than a quarter of the teachers replied that they believed their students to be well prepared to answer most; although another half believed that they were somewhat prepared. The exception was in response to a question asking students to write a story problem for a simple computation. In this case, 40 percent of the teachers replied that their students were well prepared.

In general, fourth graders were not hesitant to answer questions that may have been unfamiliar to them. In most cases, their answers showed a willingness to apply their reasoning and general knowledge to the task.

Grade Eight

The set of questions at this level contained both familiar word problems involving percentages and fractions and less common graphical problems that required students to extract and synthesize different types of information. As in the case of the fourth grade, teachers were presented with a selection of the problems and asked to respond in terms of their students preparation. It is not surprising that many teachers (between 45% and 71%) believed that their students were very well prepared to answer the word problems and those that probed students' understanding of numerical calculations. In contrast, only 20 percent responded that their students were prepared to use a map and a time/distance graph to describe a journey. Other questions—one concerning number patterns and one that asked students to synthesize and apply information from two graphs—were positively responded to by approximately a quarter of the teachers.

At this level, students' achievement reflected teaching more than in the earlier grade. Although few questions were answered well by more than 40 percent of the students, those tended to be the more familiar computational word problems. In response to these questions, high achievers on the multiple-choice section of the test were more successful than others. However, all students performed poorly on questions that asked them to observe, reason, and play with numbers. For example, there were few students who were able to draw inferences from graphical information or to identify more than one number pattern. In general, the eighth grade students were less prepared than their younger brothers and sisters to apply their reasoning in a mathematical context that was unfamiliar to them.

Grade Twelve

Again, at grade twelve, the preponderance of questions were in the area of numeration, with some questions focused on statistics and one which measured students' understanding of area. In most cases students were asked to explain or justify their response. Answers were judged on the quality of their response and credit was given for method in addition to the solution itself.

As was the case with the other grades, teachers were presented with a selection of questions and asked to comment in terms of their students' preparation. However, teachers who answered the twelfth grade questionnaire replied with reference to all grades at the high school level. With the sole exception of the value of 5!, no more than 20 percent felt that their students had been well prepared to answer any of this sample of questions. It seems that the written justification and reasoning required by these questions is not part of the usual high school curriculum.

Summary

From the teachers' point of view, students are not well prepared to answer questions that require them to synthesize information from various sources, to discover patterns, or to draw conclusions from data and argue their point of view. In general, the younger children—whether prepared or not—were more disposed to do so, and many performed well. As students progress through the grades, they appeared to be correspondingly less willing to attempt to answer questions that go beyond the specifics of the taught curriculum or required them to apply or discuss the concepts involved.

At grades 8 and 12 familiar questions that reflected the curriculum (e.g., the words problems at grade 8; questions that involved algebra at grade 12) were correlated with performance on the multiple-choice section of the test. The relationship between multiple-choice achievement and achievement with the less conventional questions was weaker, partially because few students at all proficiency levels performed well, partially because the explanations of the best students was not consistently better than those less capable.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Science Background

Currently, science curriculum is undergoing serious examination in light of recent findings indicating that United States students scored near the bottom in various international science exams. In response, studies (Science for All Americans, Research Within Reach: Science Education; 1994 NAEP Science Consensus) have generated curricular recommendations that are remarkably similar. Among those recommendations are:

Coverage should be in-depth.

By covering one topic thoroughly, students pick up skills that can be applied to other topics; they develop the habit of science. By stressing one concept and reinforcing that concept, students gain a depth of understanding than is far more substantial than a group of disparate facts. In other words, in the teaching of science, less is more.

Science should be taught as cross-discipline and multi-discipline.

Science is not a single area, but a series of related topics. Many of the concepts cross areas. The different subject areas do not exist in a vacuum and the interrelationship of the various disciplines provide a totality that can be offered in no other manner. Particularly at the secondary level, science is taught as series of discrete disciplines. Nor is science isolated from other disciplines. The relationship between science and mathematics is obvious, but science and the

humanities, the fine arts, and the social sciences are also intrinsically tied to one another.

Science education should begin early.

Elementary teachers might be tempted to give short shrift to science because they may feel they don't have the training to teach science, yet early science education is vital to developing students who have a love and respect for science. We know that despite much instruction, students often retain misconceptions particularly about the natural world. The earlier teachers can begin to delve into these misconceptions, the earlier they can start correcting them.

Science should begin with questions, not with answers.

Instead of teaching facts to students, a more engaging way of teaching might be to ask "Why do you think...?" A major characteristic of scientists is curiosity. This approach nurtures a sense of curiosity and the resulting sense of accomplishment that accompanies it. Answers should not be definitive but rather lead into more and deeper questions.

Science should be done, not read or heard.

Science is a series of discoveries based on careful observation and calibration. How better to actively engage students in science than with a hands-on approach, particularly in the early years. Students who begin with hands-on science develop facilities for observation, handling equipment and developing hypotheses that can be transferred to other subject areas. Doing science is more than following a pre-selected experiment with orderly steps leading to a predetermined conclusion; it is observing phenomena, developing a hypothesis, testing that hypothesis, and replicating that test in order to confirm the results.

Science should be cooperative.

Since science in the real world is a cooperative activity, it makes sense that science in school should be taught in the same manner. A scientist does not work in isolation, but shares goals and data. The scientific community tends to share knowledge and resources for the greater good of the scientific world, hence the world at large.

Science should have a practical application.

Unless students can assign a personal value to science, it remains arcane and distant. Using science helps students to appreciate its relevance in their own lives. Knowing science can only help consumers and voters make educated decisions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, students were asked to explain why or how something happened. In the area of earth science, they were asked to explain how water forms a canyon; in physical science, they were asked to explain condensation and evaporation. Among higher order thinking skills, students were asked to come up with a model of a natural phenomenon. They also had to complete the design of an experiment.

Teachers were given a selection of the open-ended questions prior to the test and asked to evaluate how well prepared their students were to answer the questions. At grade four, the relationship between actual performance on the questions and teachers' judgements of student preparedness for the questions varied. One-third of the teachers felt that their students were very well prepared to answer a question about evaporation and condensation, while eight percent of the students were able to discuss both evaporation and condensation successfully. On a question that asked students to consider characteristics of birds in relation to their habitats or food, 32% of the teachers felt that their students were very well prepared to answer. However, less than 15% of the students gave answers that were considered complete.

Grade Eight

The grade eight questions ranged from physics (laws of motion and friction and gravity) to physical science (the effects of temperature on a substance) to earth science (the causes of earthquakes) to scientific inquiry and process skills (steps in separating sand and sugar; reducing data in order to generalize;

generating an experiment). Other concepts included adaptation and the nature of sound waves.

As at grade four, teachers were given a small selection of questions and asked to evaluate their students' preparedness in responding to the task. Forty percent of the teachers felt that their students were very well prepared to describe an experiment, including the statement of hypothesis, design and necessary data. Approximately 15% of the students answered this question successfully. The largest discrepancy between teacher predictions and student performance was in response to the question "Explain why the moon does not fall to earth." Almost one-third of the teachers (31%) felt that their students were well prepared to answer this question, yet only 4% of the students could answer it successfully, with another 12% able to answer it with partial success.

Grade Twelve

At grade twelve, the questions focussed on interpreting data and making predictions including graphic representations. Other questions involved asking survey questions and generating a hypothesis. Students were also asked to explain why the moon doesn't fall to earth.

As in the case of the other grades, 12th grade science teachers were given the opportunity to consider the questions and respond to their students' preparedness. Given a scenario of fish dying unexpectedly, over one-fourth of the teachers predicted that their students were well prepared to generate two hypotheses and appropriate tests. About one-fourth of the students were able to generate good hypotheses and test them accordingly. Over one-third of teachers (35%) stated that their students were very well prepared to explain why the moon does not fall to earth, while 11% answered successfully. This question was problematic at both grades, with the teachers predicting much greater student preparedness than was the case. It may be that teachers felt they had covered the subject sufficiently.

At both grades eight and twelve, the majority of teachers (63% at grade 12 and 59% at grade 8) listed student indifference and lack of motivation as a very important reason why students do not perform well in science. And over 40% at both grades listed lack of preparation prior to entering the class as a cause of low achievement in science.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grades 4 & 8

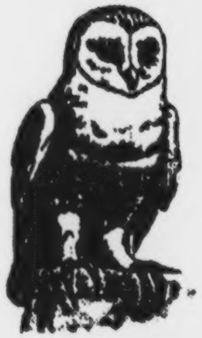
Different birds live in different areas or habitats (such as woods, lakes marshes or open areas). They eat different foods (such as fish, seeds, insects, or small animals). Birds with webbed feet swim well and live on or near water. Hummingbirds have long, thin beaks which they use to draw nectar from flowers.

Look at the pictures of the birds and notice how they are different. What do you think is the main habitat and the main food of each? Write your answers in the spaces to the right of each picture. Also explain why you chose the answer you did for each bird.

Major Concepts/Abilities Tested

Life science
Adaptation
Observation skills
Functions of characteristics
Diversity of life
Habitats

The percentage of student responses and appropriate examples are listed under each of the categories below.



Correct Answers

	Gr4	*Gr8*	
***	9%	14%	**Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.).**

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

"I think the reason this bird lives in a high branch is because he/she has sharp nails that grip into the tree and the bird's feet are shaped that way. I think the bird eats small rodents because it can swoop down and grab it with its sharp claws." (grade 4)

** 13% 25% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"This bird has long nails to catch the animals and it's beak is tilted inward to eat the animals better." (grade 8)

(Although the student mentions two characteristics, they both relate to food.)

* 6% 2% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"He eats rodents because look at the feet they were made to grip and lock. (grade 8)"

(Although there is a suggestion that the claws are for catching and grasping rodents, it has not been made very clear.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

51% 35% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"Owls live in the woods and come out at night for food and such. I read a book about on owls and found out about them. (grade 4)"

4% 4% **Mentions characteristics but does not relate.**

"This owl has very long, sharp, powerful claws and legs. The wings are powerful too. (grade 8)"

3% 5% **Relates physical characteristics not evident from picture to habitat and/or food.**

"This bird us dark and has very dull colors. It is its camouflage and when a victim comes along, the mouse probably scatters and the bird dives and catches him." (grade 8)

3% 5% **Uses "because it looks like" as a reason.**

"His nose looks like he would eat insects. To me he looks like a woods kind of bird." (grade 4)

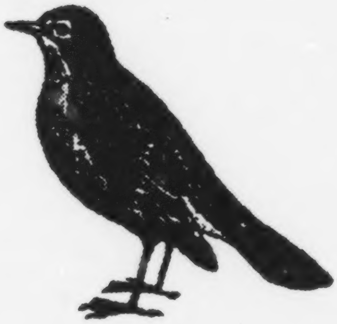
7% 4% **Other incorrect response.**

"He eats caterpillars because it shows him sitting on a tree and trees usually grow in the woods and caterpillars crawl on trees." (grade 4)

3% 8% **Blank**

Comments/Owl

It's clear from the results that students learn a lot about owls and feel compelled to share their knowledge. The owl is the only bird of the four given that elicited such a large number of students responding in this way. It could be that the owl is so interesting and unique and such a common cultural symbol that students know more discrete facts about them. It is possible that students could have incorporated prior knowledge into a correct response. Many students did. For example, if a student knows that owls eat mice, then he or she may determine that the claws are large for tearing mice simply from knowing what owls eat and observing the picture of the owl. Fewer students left this bird blank, indicating that they have some knowledge about owls. However, it seems as though those blanks were picked up in the prior knowledge category.



Correct Answers

Gr4 Gr8

*** 5% 13% **Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.)**

"The beak looks like it digs worms from the ground because of the pointed end and the feet are somewhat webbed and looks like it could walk on mud of marshes." (grade 8)

(Although the bird did not have webbed feet, the picture could been interpreted as such. The student received full credit for the response because he/she was able to make the appropriate inferences. We were not looking for correctness of response regarding what the actual bird ate and where it lived; we were assessing students' ability to observe a characteristic and make a judgement based on that characteristic.)

** 17% 24% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"They have pointed beaks to get seeds from the ground." (grade 4)

* 9% 3% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"This bird has webbed feet and there are many insects near marshes." (grade 8)

(There is no explicit relationship between webbed feet and insects; the inference being that a bird with webbed feet would live near the water, and because insects live near the water the bird would eat insects.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

28% 21% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"This bird is a robin and usually found in the wide open area; for example, your yard." (grade 8)

5% 6% **Mentions characteristics but does not relate.**

"He doesn't have claws or a short fat beak. Instead it has a long skinny beak with webbed feet." (grade 4)

8% 9% **Relates physical characteristics not evident from picture to habitat and/or food.**

"I'd say it's main habitat is in a tree because this is a rather small bird and needs a warm and safe place for itself and its babies. I'd say it's main food is worms or insects because it is a small bird and would not be able to kill or swallow mice like an owl." (grade 4)

(Although it may be prior knowledge that leads the student to believe that the bird is small, the picture does not indicate the bird's size. In the pictures, the largest bird, the heron, actually appears to be the smallest. When dealing with size, there are some size qualities that we do accept, for example, students' responses mentioning the heron's long beak are accepted as correct because, in relation to the rest of the heron's body, the beak is very long.)

8% 8% **Uses "because it looks like" as a reason.**

"This bird looks like it could be seen anywhere and live comfortably at any of the locations. It also looks like it eats worms and other insects." (grade 8)

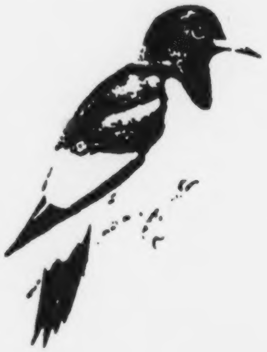
13% 5% **Other incorrect response.**

"This bird has eyes on each side of his head so he or she can see things coming toward him. That is for protection." (grade 4)

7% 11% **Blank**

Comments/Robin

About the same number of students answered this question correctly as did the previous question. Although the robin is as familiar to most students as the owl, fewer students answered in the category of prior knowledge. It could be that, without seeing the colors of the birds, students did not recognize the robin. There may have been some confusion for some students who recognized the robin, but then saw the webbed feet and could not adjust their prior knowledge of robins to what they saw in the picture.



Correct Answers

Gr4 Gr8

*** 8% 17% **Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.)**

"It probably lives in tree trunks because it can dig its claws into the side of the tree to build its nest. Also, I think it eats bugs because it has a long beak that it can put into a tree and pull out bugs." (grade 4)

** 17% 24% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"This bird's long beak looks good for reaching into the home of insects because it's long and narrow." (grade 4)

* 7% 3% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"The woodpecker likes to carve out wood from the trees with its sharp beak and finds and eats the insects." (grade 8)

(The student response is fairly clear, but the relationship between the characteristic and the food is very vague, bordering on incorrect.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

29% 20% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"The woodpecker pecks on trees in the woods to get insects out of the trees."

(This answer is very similar to the third response above, but there is no mention of the beak or its qualities.)

8% 6% **Mentions characteristics but does not relate.**

"He has sharp claws and a pointed beak." (grade 8)

7% 5% **Relates physical characteristics not evident from picture to habitat and/or food.**

"Since I have not seen a bird like this around, I suspect it hides itself in someplace where it can't be seen easily—the woods. It is a smallish bird so I would think it eats insects." (grade 8)

7% 8% **Uses "because it looks like" as a reason.**

"He looks like he is too small to eat meat." (grade 4)

10% 5% **Other incorrect response.**

"This bird looks for a place where he can see more clearly." (grade 8)

8% 13% **Blank**

Comments/Woodpecker

Student responses here are very similar to those of the robin. It appears that the more usual looking the bird, the less students have to say about them. Compare the results of the robin and woodpecker with those of the heron which follow.

Correct Answers

Gr4 Gr8

*** 14% 34% **Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.)**

"The long legs can hold him above the marshes and the long pointed beak is ideal for stabbing fish." (grade 8)

** 20% 22% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"This bird has long legs, a long neck and a long beak that could allow him to walk in the water and dunk his head into the water to get fish." (grade 8)

(Although three separate characteristics are mentioned, the student only discusses them in relation to food.)

* 7% 3% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"They have webbed feet and if they live near or on water then they are most likely to eat fish." (grade 4)

(Here there is an inference of the relationship between webbed feet and fish as food although it is not well elaborated.)

"Because he has long legs you would probably see him near the water so he would probably eat fish." (grade 8)

(There is no indication as to why long legs would make the bird a water bird.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

29% 16% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"It's a flamingo and flamingoes live in marshlands and marshlands are mostly bugs and fish." (grade 4)

8% 5% **Mentions characteristics but does not relate.**

"It has long webbed feet, a long, sharp beak and a long neck." (grade 8)

5% 1% **Relates physical characteristics not evident from picture to habitat and/or food.**

"He is big and is easily able to hunt small animals." (grade 8)

5% 2% Uses "because it looks like" as a reason.

"This bird looks like a bird from the jungle who eats fish." (grade 4)

6% 4% Other incorrect response.

"This is the ancestor of the swan." (grade 8)

8% 13% Blank

Comments/Heron

At both grade levels, students were most successful when writing about the heron. Approximately ten percent more of the fourth graders and twenty percent more of the eighth graders were able to answer this question successfully than were able to answer the other birds given; owl, robin and woodpecker. What is it then about the heron that lends itself to better responses? One answer may be that with its long legs, neck and beak, the characteristics are far more exaggerated so students are able to more easily make connections about adaptations. Another might be that the bird is far more exotic and therefore less familiar to New England students. The owl carries cultural baggage, the robin and woodpeckers are both fairly mundane, but the heron is not an everyday bird.

Comments

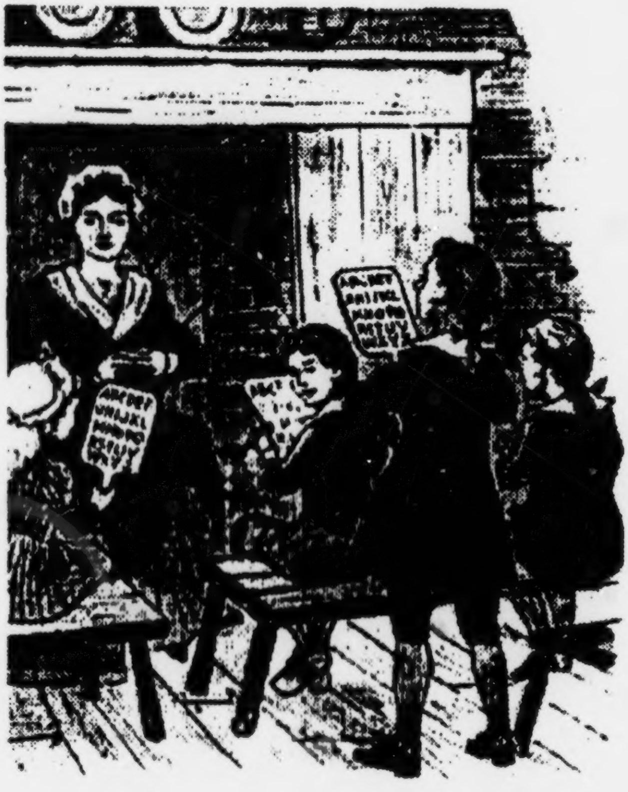
The responses to this question suggest that although many students are able to make inferences based on observations, a large number have problems with this type of thinking. It is a habit of mind rather than a piece of discrete information. One way of opening students to this type of thinking would be to begin with a question rather than with a fact. For example, these questions, "Why does the owl have such strong, sharp claws?" and "How do the owls' claws affect how they survive?" open the door to a more unified and systematic way of thinking and looking at something. By connecting the structure of the animal with its environment and diet, we can show the interrelatedness of all species. Students will eventually come around to the desired conclusion, but it will be through a series of observations and discoveries rather than memorization and recitation. This type of subject leads to a discussion of adaptation and how, over the generations, species carried characteristics that enabled them to survive.

Another way of getting students to examine function and form is to ask them to draw a bird that lives in the water and eats fish, as opposed to a bird that lives near the water and eats plankton. This method would direct students to forms of adaptation.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 4



The picture above shows children living in the early American colonies. How were the day-to-day lives of these children like your life? (Discuss many ways that they were probably the same.)

Major Concepts/Abilities Tested

Sense of history

Relationship with the past

Specifics of colonial life

The percentage of student responses and appropriate examples are listed under each of the categories below.

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

Correct Responses

*** 11% **Evidence of abstract concept(s) i.e., makes generalizations about human nature; discuss "big ideas" with or without comparative statements; commonalities within cultures (work ethic, love, human needs); technology vs. environment.**

"Both now and then we have arithmetic and reading. We also have a teacher. Also, like then, there is always a genius in the class and there's always someone everyone picks on. Then and now we use the alphabet."

(Initially this response seems to be a list of what similarities are made in the school, but the student jumps from the simple components of a classroom to a much larger issue—one of human nature. The student recognizes something essentially human about the condition of being the genius or being the goat in a classroom.)

** 21% **Literal, more concrete response—includes two or more comparative statements and elaborates beyond specifics in the picture.**

"We both live in homes. We both go to school. We both are humans. We both are boys. We have teachers."

(This response also recognizes the human condition, but that recognition seems a physical rather than an emotional or spiritual connection.)

* 39% **List of a range of similar items rather than comparative statements.**

"They went to school. They had clothes. They ate everyday. They did chores. They lived in houses."

* 23% **List of similar items based predominantly on the picture.**

"They sit in chairs like we do. They learn the alphabet like we do. They braid there hair or put them in pony tails like we do."

*** *very good response*
** *good response*
* *minimally correct response*

Incorrect Responses

3% **Limited response i.e., only one similar item or comparison given.**

"Their life in the early American colonies wasn't much alike, but we had to go to school."

2% **Incorrect/inaccurate response only.**

"No desk, no T.V., no books, no chalkboards, no erasers, no clocks, no lights, no crayons, no sinks, no coat closets."

1% **Blank**

Comments

In the elementary grades the study of history should develop in students a sense of the significance of the past and a unity with those whose lives they study. By the fourth grade, students seem to have developed a sense of similarities and universalities between then and now. The task is to instill a sense of importance to the details that they recognize. Is the fact that they also use chairs an important bridge in the chronological gap? Students can begin to assign values to the details they note. By making history real, students can begin to appreciate the significance of the past on the present. That 94% of the students were able to note some similarities indicates that students are on their way to achieving this understanding.

How were the day-to-day lives of these children different from yours? (Discuss many ways that they were probably different.)

Major Concepts/Abilities Tested

Sense of history

Relationship with the past

Specifics of colonial life

Correct Responses

*** 5% **Evidence of abstract concept(s) i.e., makes generalizations about human nature; discuss "big ideas" with or without comparative statements; commonalities within cultures (work ethic, love, human needs); technology vs. environment.**

"We have lunch and recess. We get to watch videocassettes. We have other activities such as art, music, gym, media, and some advanced classes. We have a lot of new technology. They didn't have those things, but they had some things that we don't have such as: rich soil, good farming, different clothes, and more and more nature and natural resources."

** 23% **Literal, more concrete response—includes two or more comparative statements and elaborates beyond specifics in the picture.**

"Some of the ways we're different were their houses are colder than ours. We have machines to do most of our work. We have flowing water and electricity. We have bigger schools and telephones."

* 44% **List of a range of different items rather than comparative statements.**

"I have ridden in a car and plane. I have a desk. I have a T.V. I have a microwave. I have a radio. I have a nintendo. I have electricity."

(Although the student implies that the colonial students have not done what he has, there is no elaboration of this. If the student had stated, "I've ridden in a plane, and the colonial student probably walks or goes places in a horse drawn wagon," then this statement would have been considered in the second category of responses.)

* 24% **List of different items based predominantly on the picture.**

"They are wearing some clothes that we don't. The boys have to wear bows in the back of their tails, but us boys don't. The schools in the early American colonies are more different than the schools we have now. They dress up like pilgrims."

"They wrote on slates instead of on paper. They sat on wooden benches instead of metal chairs. They didn't have desks like we do."

*** *very good response*
** *good response*
* *minimally correct response*

Incorrect Responses

3% **Limited response i.e., only one different item or comparison given.**

"They wore different clothes and the men wore wigs."

1% Incorrect/inaccurate response only.

"We go to a school not a house and we don't wear uniforms."

"They stay home and go to school."

0% Blank

Comments

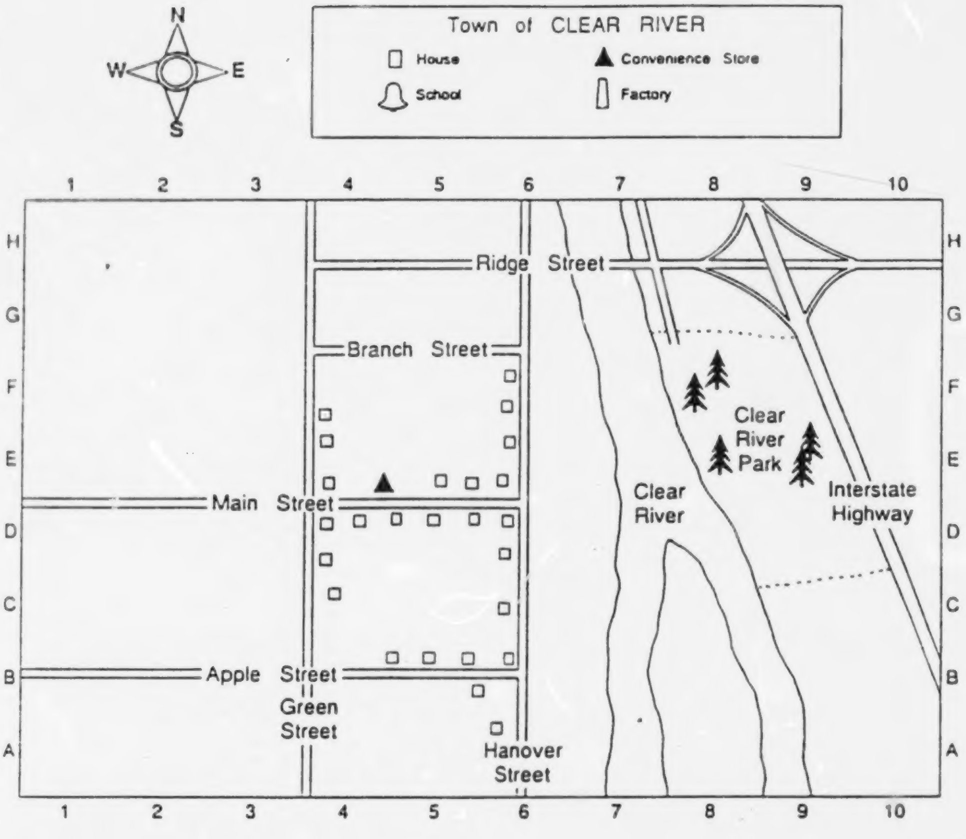
Students performed extremely well on this question with 96% of the students able to provide at least an adequate response. As in their selection of similarities, students, by their choice of differences, are indicating what they feel is important. Those differences students noted (mainly technological) signify vast differences in our culture and how we live. Students appear to be able to note and appreciate these differences while maintaining a unity with the past.

These two questions direct students to make connections between their own worlds (starting with the classroom and expanding outward) and those of the past. By asking what is the same and what is different, teachers begin to encourage a habit of mind that enhances understanding of today's world.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 4


N
W
E
S
Town of CLEAR RIVER
House
Convenience Store
School
Factory
1 2 3 4 5 6 7 8 9 10
H G F E D C B A
Ridge Street
Branch Street
Main Street
Apple Street
Green Street
Hanover Street
Clear River
Clear River Park
Interstate Highway

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

A school and a factory are going to be built in Clear River. Where do you think the best locations for these would be? Draw a bell 🔔 on the map where you think the school should be. Draw a smokestack ▯ on the map where you think the factory should be. In the proper places on the next page, give the map coordinates of the school and the factory, and explain why you think the locations you chose are the best locations.

Major Concepts/Abilities Tested

Process skills

Map skills

Human geography

Graphic representations

Part 1—Placement on Map

	School	*Factory*	
*	90%	90%	Accurate placement on the map
	4%	4%	Inaccurate placement in river
	2%	1%	Other inaccurate placement
	3%	3%	Blank

* *correct answer*

(Any placement in the street, park or median of the highway was considered inaccurate.)

Part 2—Give Coordinates

Correct Answers

	School	*Factory*	
**	55%	57%	Correct coordinates using letter and number
**	3%	2%	Correct coordinates using other description for example, the corner of Green and Apple Streets
*	10%	9%	Partial correct coordinate (letter only, number only or partial description)

Incorrect Answers

School	Factory	
3%	3%	Coordinates mismatched to placement on map
4%	4%	Evidence of misunderstanding of term "coordinates" (used as a verb)
6%	6%	Other incorrect response
19%	19%	Blank

** *good answer*
* *minimally correct answer*

Part 3—Explanation of Selected Site

Correct Answers

School / Factory

*** 14% 16% **Two or more acceptable reasons with elaborations or one reason with two acceptable explanations**

"I put the school on Branch Street because I thought that there is a nice big spot to play in and it is near most of the houses so kids won't have far to walk."

"E-1 is pretty far from Clear River and Clear River Park. The factory would not pollute it very much. Also, the workers wouldn't have too far to drive."

(Reason and explanation may seem redundant, but the explanation is an elaboration of the reason. In other words, in the first example, the reason the school was placed on Branch Street is that it has a "nice big spot" and the explanation for that reason is "for kids to play in." In the second example, there is one reason, the factory would be far from Clear River (meaning the river and not the town) and Clear River Park, the explanations are that the river and park wouldn't get polluted and the workers would have far to drive. The factory was situated in coordinates E-1.)

** 52% 52% **One acceptable reason with one acceptable explanation**

"I think that's the best place for the school because it's near the houses and people will not have to go too far to get to school."

"I put the school there because there are a lot of houses and they have to walk far and that will say all their energy and plus if you forget your homework you don't have to walk so far back."

(Although this student has three explanations why the school was placed near the homes, the three explanations are really the same that is that students would not have so far to walk.)

** 8% 7% **Two or more acceptable reasons without explanations or explanations without reasons**

"School—It is close to the houses and there is a large area."

* 17% 16% **One acceptable reason without explanation or explanation without reason**

"I put it there because it is close to Clear River Park."

(Here there is a reason for placing it but there is no explanation as to what being "close to Clear River Park" means.)

*** *very good response*
** *acceptable response*
* *minimally acceptable response*

Incorrect Answers

School / Factory

5% 9% **No acceptable reasons or explanations or other incorrect answer**

"The school coordinates a library for people to take books out or for people who have a hard time working so they can do research."

"The factory coordinates a cleaning product so it will be easy for people to clean things and the people who need the product will shop at the factory."

(Obviously the student who gave the above responses misunderstood the meaning of the word coordinates and that misunderstanding influenced the rest of the answer. Four percent of the students answered in this manner.)

3% 0% **Blank**

Part 4—Environmental Issues

School / Factory

4% 49% **Evidence that pollution/ environmental issues were considered in selection of site**

94% 51% **No evidence that pollution/ environmental issues were considered in the selection of site**

(Environmental concerns were noted but not scored as correct or incorrect.)

Comments

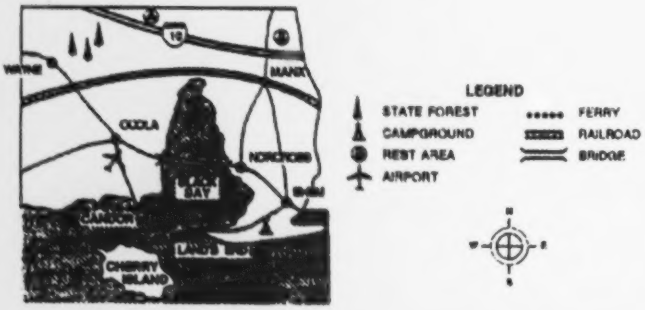
Generally fourth graders do not leave questions unanswered. That 19% of the students left the coordinates portion of this question blank indicates a possible problem with the vocabulary or a mismatch with fourth grade curriculum. Since this question was a part of the larger question, it may be that students felt more comfortable with leaving a portion of the question blank.

This question focuses not only on the process skills involved with map reading and interpretation of visual data, but also on the concept of community. In order to answer the explanation portion of the question, students had to consider the community and its needs, whether economic, social or environmental. They had to analyze the human interaction with the environment in order to reason why they placed something where they did. This concept of community can be enlarged to the idea of global community. The world has been made smaller through electronics, and students are aware of events taking place across the globe. At this point, they can start making the connections that tie us to the rest of the world.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 4



You and your family live on Cherry Island. (See the map above.) Some friends are coming for a visit. They will be travelling east on Route 10. Write a set of directions telling your friends how to get to Cherry Island.

Major Concepts/Abilities Tested

Process skills

Map skills including ability to read and use maps

Ability to organize information

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** 13% **Correct route including all stages using cardinal directions.**

"Go south past Manx. Go southwest to Norcross. Go west past Black Bay. Go south to Candor to take a boat across."

(Although this response appears brief, it covers each stage adequately and correctly.)

** 12% **Correct route including all stages using left/right directions.**

"Go down route 10 until you meet a four-way with a rest area, take a right. Follow that road until you reach a railroad track that goes through the town of Manx. Take the right hand road. Follow that road until

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

you find another town, go right across the bridge, keep going straight until there's a three-way, go left all the way down that road to a ferry. Take the ferry to us."

*** 6% Generally correct route—confuses cardinal or left/right directions.**

"First you will pass the State Forest (on your left). Then a little before a rest area, go down to Manx until you come down to a railroad. Now there is a fork in the road. You will take a left one now you come to a bridge. Go over the bridge, you see an airport. One your side of the airport, go down that road. Now you are at a ferry. Go down the ferry and you car here."

(This student confused left and right but otherwise answered correctly.)

*** 27% Generally correct route—one or two stages confused, incorrect or missing.**

"To go to my house travel up route 10. When you see a road on your right, turn right and then you will end up in Manx. Keep going down the road that you did on route 10. You'll end up in Norcross. Go over the bridge. Keep on going. You will see a road going south. Go down it to the Ferry, that will take you to Cherry Island."

(This response does not mention that the road forks in Norcross.)

** *good answer*
* *minimally correct answer*

Incorrect Answers

8% Gives directions (cardinal or left/right with little or no reference points indicated.

"1. When you come to a rest area, stop and turn right. 2. When you turn, keep going straight until you come to the bridge. 3. When you cross the bridge turn left and go straight. 4. When you come to the ferry, take it across to Cherry Island."

15% Directions based on inaccessible routes or transportation other than car or ferry.

"As you are traveling east on route 10, stop at the rest area. Then start south and you should come across a railroad. Take a train west then south until you reach Ocola. Take a plane south until you reach Candor. Finally take a ferry to Cherry Island."

15% Other incorrect response.

"Start at Shem. Go straight. Go over the bridge. Keep on going then take a left. Go straight again."

4% Blank

Comments

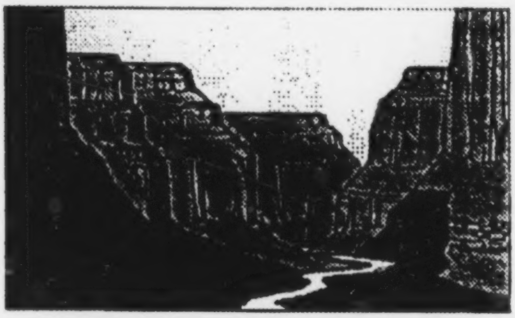
Over half the students were able to meet the requirements of this question with some success, but half of those missed a step or confused directions. This question measured two separate skills: the first is the ability to comprehend and interpret a map correctly; the second is the ability to articulate the process and present it clearly and logically. Most fourth graders are accustomed to finding something on a map, but far fewer are able to describe the correct way of getting there. This requires that students organize the material in a logical manner and present it with the appropriate details to make it recognizable. The need for this type of writing crosses all other subject areas, especially science and math, where clearly articulated sets of procedures are necessary for duplication of results.

Probably most fourth graders could point out Cherry Island on the map, but we see that they have some difficulty in describing how to get from one place to another. This type of question illustrates the need to do rather than hear this type of social studies.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 4



People claim that the canyon shown in the picture above was created by water. Explain how water could do this.

Major Concepts/Abilities Tested

Earth science

Formation of land

Changes in the earth's surface

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 16% **Discusses erosion with accurate elaboration (e.g. moving water, long time, etc.).**

"A long time ago, the land had lots of rivers running though it, and it began to wear out. It caused erosions and after many years, it formed this."

** 22% **Erosion; some elaboration, no inaccuracies.**

"Water could cause erosion. Erosion is the wearing away of the earth's surface by wind or water."

(Here the student does not mention the length of time involved. "Wearing away" might be construed as implying a length of time, but we did not feel we could infer that from this answer.)

"Water could have created the canyon shown above because if you think about it,

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

if you have a lot of sand or dirt, then rushing water comes through, sooner or later some of the sand will disappear."

(This answer does address the time issue to some extent, but implies that it is the strong force of water on loose soil that causes erosion rather than the constant flow of water on packed soil.)

* 21% **Erosion; no elaboration or some inaccuracies.**

"The force and power of the water can break off pieces of the earth. Then after many days it forms a canyon."

*** *very good response*
* *good response*
* *minimally correct response*

Incorrect Answers

6% **Rain washes away rock.**

"Water could have done this by raining a lot. Say there was a rainstorm, if the rain came down real fast it could knock off pieces of rock and it would chip more off every time it rained there.

6% **Sand/material is deposited to form walls/mountains.**

There might have been a big erosion a long time ago and it probably carried the rock and stone in it to this place and all the rock and stone piled up."

6% **Answer implying that the water has drained away, leaving the canyon i.e., there used to be a lot more water.**

"I think that there was a flood and that water came running down the canyon and the water dried up and it formed the shape of the canyons.

7% **Discusses other natural phenomenon e.g., earthquakes, glaciers, volcanoes.**

The canyons were once part of all land. One day and earthquake took place which cracked the land, Water came in and the pressure of it formed the canyon."

16% **Other incorrect response.**

2% **Blank**

Comments

About 60% of the students were able to answer with some degree of correctness. Students demonstrated many misconceptions of how water forms a canyon: water deposits on the rock to form the canyon walls; the force of water rather than its steady flow causes the canyon; the canyon was formed instantaneously as the result of an earthquake or flood. The idea of great sudden violence seems to override the idea of gradually wearing away the surface.

Twenty percent answered "erosion," but did not define it or defined it incorrectly. Less emphasis on the word erosion and more emphasis on the concept it represents would allow students to articulate the concept in their own vocabularies. If students are clear on the concept, then the introduction of the scientific vocabulary can only enhance their understanding; if they are not clear, then the introduction of that vocabulary tends to mask misunderstanding with an appropriate word.

Many students have a problem with the concept of time, especially when trying to define a long period of time. They are not able to distinguish between one hundred years, one thousand years, and one million years. So students who define erosion as running water wearing down rock over a long period of time may not be entirely correct, depending on their definition of a long time.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 4

Every time Tom boils water to make spaghetti, the insides of his kitchen windows get wet. Explain how this happens.

Major Concepts/Abilities Tested

Physical science

States of matter

Transformation of water

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 8% **Discusses water EVAPORATING and CONDENSING (may or may not actually use terms); explanation clear and completely correct.**

"The water boils and turns to steam. Steam is a gas called water vapor. Water vapor gets on the windows. The windows are cold so the water cools off causing the water to condense or change into little droplets of water which make the windows wet."

** 13% **Discusses water EVAPORATING and CONDENSING (with or without using terms); explanation somewhat vague or lacking; may contain some inaccurate statements.**

"Before the water boils, it's just water. But when it boils it makes steam. The steam is made by little drops of water. When the steam cools, the water drops go back to water and makes the windows wet because the steam was in the air."

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

(Although the student displayed some correct notions with the mention of heating and cooling, the student does not define steam as a gas, but rather as smaller drops of water. There is not enough sense of transformation from liquid to gas and back to liquid states to merit the first classification.)

* 54% **Discusses ONLY EVAPORATION or production of STEAM and not transformation back into liquid.**

"The steam rises up and covers the window with water because the steam cannot get out."

"When you boil water for a period of time, steam starts to rise. When the steam rises and its near a window, the window will start to get damp."

"Every time Tom boils water to make spaghetti, the insides of his kitchen windows get wet. This is because the water evaporates and clings to the window."

Students whose answers fell in this category appear to have a general sense of the relationship between the boiling water and wet window, but don't understand the transition so they simply say the steam sticks to the window.

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

11% **Suggests that the HEAT is creating the moisture.**

"Because the steam from the water is so hot when it goes up in the air it brings water along and it hits the window and the window becomes wet."

15% **Other incorrect response.**

1% **Blank**

Comments

This question attempts to find out what children really understand about the transformation of matter. Can they do more than offer the terms *evaporation* and *condensation* and explain that although matter changes state, it remains the same substance?

We did not expect students to explain this on a molecular level although a few did attempt it successfully. More than half were able to adequately explain evaporation, yet only 21% were able to explain evaporation and condensation. Although students appear to be aware that steam is a gaseous state of water, they do not seem to understand the converse, that steam can revert back to water. Some described steam not as a gas but as tiny droplets of water floating through the air. Often students knew the terms *evaporation* and *condensation*, but were not able to elaborate.

Student responses to this question point to two potential difficulties that are not unrelated. The first is that students are not clear on the concept of transformation of matter. At the early grades, this is the type of science that students can easily understand if it is taught in a concrete manner; i.e. if students are allowed to do experiments, make observations, and draw conclusions about the states of water.

The second problem is that students use of the correct word for a concept may mask their misconceptions. This can happen when the technical vocabulary is introduced before the student has some understanding of the concept. If a teacher accepts the answer "evaporation" as an explanation of what happens to water when it is heated without learning what evaporation means to that student, then it is possible that the student and teacher mean different things. The longer these types of misconceptions go uncorrected, the more difficult it becomes to amend them. Some of the current research about prior knowledge indicates that a demonstration of the concept along with a reading in a text is successful in correcting misconceptions.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 4

How could you show a classmate how water can create a canyon?

Major Concepts/Abilities Tested

Earth science
Concept of erosion
Modelling
Reasonableness of response

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 5% **Subjects dirt, sand, rock or other material to MOVING water passing through the material over a PERIOD OF TIME.**

"I would get a big long plastic bowl or something like that and stack dirt up to the sides and run water through it constantly. After a long period of time, check it and see what happens. Hopefully it would work out right."

"You could take a rough rock and somehow get water to run onto it without wasting *too* much water until it gradually got smoother and smoother."

** 11% **Subjects dirt/sand to MOVING water (accurate simulation of the effects of erosion—not necessarily over a long period of time).**

"You can take a jug of water, put piles of dirt on a platform and pour the water on the platform and it will make like a tunnel in between the piles of dirt."

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

(The difference between the first two correct answer categories is that the first accounts for some length of time to demonstrate erosion and the second uses a material such as sand that is not compacted, so it will show the effects of erosion without a repeated stream of water. Both were considered to be accurate.)

** 16% **Subjects material to moving water—not as good a simulation as above.**

"I would take loose sand and pour water through it. The water would make small canyons wherever it went."

(The first three responses all suggest that the water flows through the material.)

* 12% **Instructional devices (static model, pictures, report, video to illustrate rather than to demonstrate).**

"You could show him pictures of one mountain over the years. You could show him how water erodes the mountain each year."

*** *very good response*
* *good response*
* *minimally correct response*

Incorrect Answers

(Generally, incorrect answers did not consider the flow of water.)

20% **Mixes dirt/sand/rock with water implies standing water when poured and/or compacting of the dirt.**

"If you took a bowl, soil and water you could fill the whole bowl full of soil and then pour water down the middle. In the middle it wouldn't be so much soil. It would have been pushed to the sides."

7% **Response implies dissolving something in water.**

"By getting a bucket of water and some paper crumpled up with a rock in it. Put the paper in the water and it will start to rip."

26% **Other incorrect response.**

4% **Blank**

Comments

Many students (20%) did not take into account the idea of a flow of water wearing away the surface. They tended to suggest a more stagnant representation such as putting some sand in a bucket and running water into the bucket. In order to be perfectly accurate, there had to be a means of escape for the water.

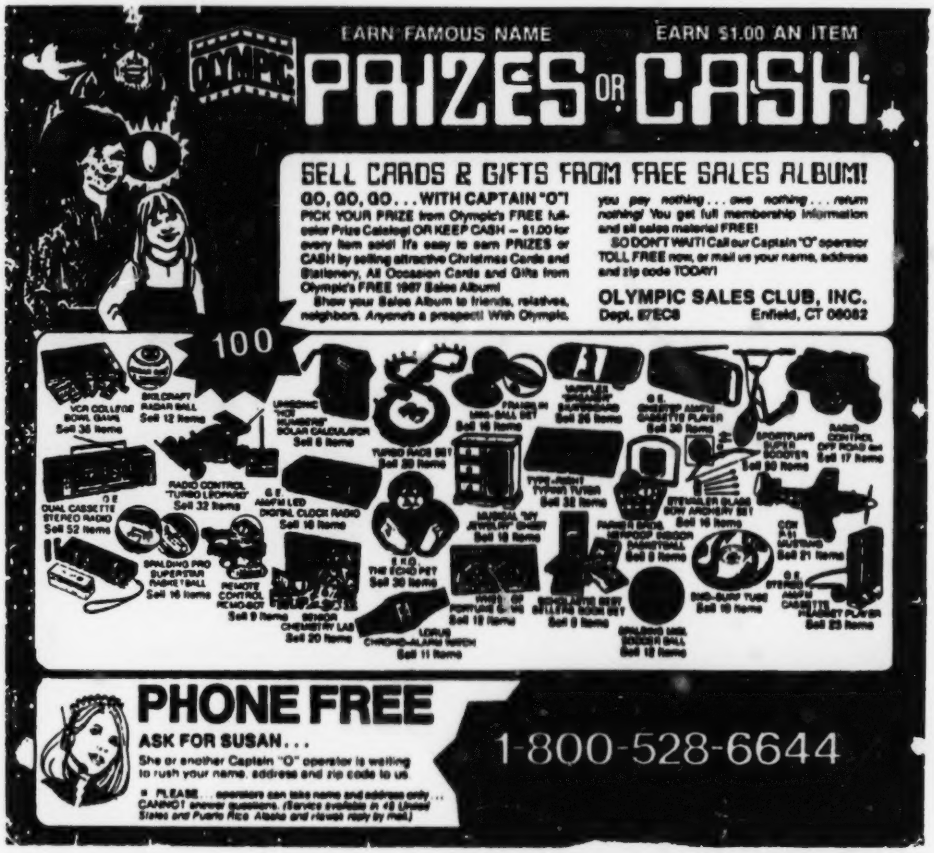
A model makes the abstract concrete and the inaccessible accessible. Students' models can indicate their levels of understanding. Those who may give a correct written description of erosion may show their basic lack of understanding of the concept by suggesting that water be placed in a bucket. Some students suggested filming erosion, indicating a weak concept of time; others suggested visiting the same place year after year, indicating a sense of time, but not the idea of reasonableness of response.

This type of question lends itself very easily to becoming an in-class experiment. With students working cooperatively, the combined knowledge and skills would probably generate better results than those shown here.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 4



EARN FAMOUS NAME EARN $1.00 AN ITEM

OLYMPIC

PRIZES OR CASH

SELL CARDS & GIFTS FROM FREE SALES ALBUM!

GO, GO, GO . . . WITH CAPTAIN "O"! PICK YOUR PRIZE from Olympic's FREE full-color Prize Catalog! OR KEEP CASH — $1.00 for every item sold! It's easy to earn PRIZES or CASH by selling attractive Christmas Cards and Stationery, All Occasion Cards and Gifts from Olympic's FREE 1987 Sales Album!

Show your Sales Album to friends, relatives, neighbors. Anyone's a prospect! With Olympic, you pay nothing . . . owe nothing . . . return nothing! You get full membership information and all sales material FREE!

SO DON'T WAIT! Call our Captain "O" operator TOLL FREE now, or mail us your name, address and zip code TODAY!

OLYMPIC SALES CLUB, INC.
Dept. 87EC8 Enfield, CT 06082

100

PHONE FREE

ASK FOR SUSAN . . .

She or another Captain "O" operator is waiting to rush your name, address and zip code to us.

* PLEASE . . . operators can take name and address only . . . CANNOT answer questions. (Service available in 48 United States and Puerto Rico. Alaska and Hawaii reply by mail.)

1-800-528-6644

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

BEST COPY AVAILABLE

Explain in your own words all the things you have to do to get prizes or cash.

Major Concepts/Abilities Tested

Reading for detail

Drawing inferences

Correct Answers

*** 22% **Answer communicates the basic idea that you earn cash or prizes by selling items and refers to procedure for doing this.**

"Call Susan for a catalogue. Then you have to sell items. You can get different prizes depending on how many items you well. You can get cash because if you want you can have a dollar for every item sold."

** 40% **Answer generally correct, but is vague concerning procedure.**

"You get prizes or cash if you sell a number of items from their company. Example: I sold 30 items, I get a boom box."

* 7% **As above, but misinterprets or oversimplifies terms of the offer.**

"You sell cards, gifts and stationary, and then you can either keep the cash or send it in for prizes."

*** *complete response*
** *adequate response*
* *minimal response*

Incorrect Answers

8% **Answer misinterprets advertisement.**

"You can sell items, gamble, try to win on a T.V. game show. You can also save up money you earn to buy gifts."

"Call 1-800-528-6644 and ask for Susan. She'll give you details. Then pick your prize or money and mail name and zip-code."

15% **Answer depends on inappropriate use of prior knowledge rather than reading the text.**

"If you want cash or prizes you have to enter a contest or suppose you [illegible] the one millionth customer in a store, you could win prizes or cash ... You could also raffle things off by writing the amount of cash or prizes on a piece of paper and there're numbers on everybodys piece of paper that they have and whoever get the number gets the prize or amount of money."

7% **Other incorrect.**

"You have to sell 35 books, 50 candybars, and 5 baby kittens and then call 1-800-538-6644 and ask for Susan and tell her you've sold everything."

2% **Blank**

Comments

Whether we desire it or not, advertising forms an important influence on all our lives. The stimulus for this question was an actual advertisement that appeared in a children's magazine. The question itself investigated how well the students were able to infer the terms of the offer and to identify the procedure required. This was not an easy assignment. The layout was cluttered and deceptive, the prose had to be read carefully for details, and large inferences had to be made. The "cash reward" led many students to believe that they could keep all the cash from their sales. Others did not realize that "free" referred only to the sales catalogue, not to the items themselves. The terms of the offer (selling cards and gifts) were carefully couched in enticing prose dwelling on the ease with which cards could be sold. Alternately, the requirements could be inferred from the descriptions under the picture of each "prize." All in all, this was a realistic test of students' ability to sort out the facts from a set of half-truths and deceptive prose.

Advertisements use many ways to convince people to do something. What are some ways that this advertisement tries to convince you? Give examples from the advertisement.

Major Abilities Tested

Relating information to prior knowledge

Applying critical evaluation

Correct Answers

*** 22% **Answer refers to at least three correct persuasive techniques.**

"This advertisement tries in one way to convince you by enlarging the good things you get. Another way is by not giving that much information about the things you have to do. They write words that make kids want to try to win the prizes and cash. They make it sound very easy and fun."

"They say you pay nothing, owe nothing, return nothing, but you have to sell things. They also say you get everything free, but you still have to sell something. They show picture of great prizes but they're probably cheap!"

** 24% **Answer gives examples from advertisement but does not explain their use.**

"The advertisement says you get everything free, even the phone call is free! Earn one dollar for every item sold!"

"Earn famous name! Pick your own prizes! or take cash! You pay nothing, owe nothing, and return nothing!"

"All you have to do is sell gifts and cards and you get free prizes. And the phone call is free and you get $1.00 per card or gift you sell."

"They show prizes that make you want to sell stuff. And they show happy children. That makes you want to sell and look happy too."

*** 27% Answer refers to a single technique or gives a general response. Some lack of understanding.**

"The ad tries to convince you by showing you all the things you could win."

"They convince you that you will get money and prizes if you sell something but you might not."

*** *complete response*
** *adequate response*
* *minimal response*

Incorrect Answers

4% Answer misinterprets advertisement.

"They tell all the good things about the item. They put famous people on T.V. trying to sell an item to you and a lot of other things too."

9% Answer refers to unrelated knowledge or gives general discussion of advertising.

"Suppose you listen to the radio say something about hair tonic and they convince by giving people money just to see it work, just for people to buy their products. But when you try, it does something wrong instead of something good. And they also give you commurcules on body oil, but when you try it yourself your body gets kind of pale, and you can't take it off."

12% Other incorrect response.

"With Olympic you pay nothing, owe nothing, return nothing."

2% Blank

Comments

In response to this question, students were asked to assume a critical stance. Assuming that they had all had some experience with advertising techniques, they were asked to identify their specific use in the advertisement. This allowed them to step outside the text itself and to judge the way in which the author uses words and pictures to convince the reader. This question calls for a recognition of the emotive power of words to persuade and confuse. Although most students were able to recognize examples of one or more specific techniques that were used in the text, only a few were able to characterize those techniques and describe their effect.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 4

A very short story is going to be broken down into four parts. Read the first part; then answer the question that follows it. Only after you have answered the question that follows one part are you to go on to the next part.

Part 1

The stillness of the morning air was broken. We headed down the bay.

What are some questions you would want to have answered so that you would understand what this story is about?

Major Concepts/Abilities Tested

The ability to recognize the kinds of information needed to construct meaning from the text.

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** 40% **At least two questions that reflect experience with narratives, e.g., questions about plot, characters, setting, title, genre.**

"What happened? Why did you go down to the bay? Was someone screaming? Was there a fire, tornado, or hurricanes?"

"What month is it? What bay are they going to? Why are they going to the bay? Who is we? How was the stillness of the air broken?"

* 15% **At least one question described above.**

* 10% **Other legitimate questions.**

"What does 'headed down the bay' mean?"

"How was the morning air broken?"

* 22% **Some combination of above.**

** *very good response*
* *adequate response*

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

Incorrect Answers

5% Irrelevant questions or questions about insignificant details.

"Was the bay pretty?"

7% Other incorrect.

1% Blank

Comments

The purpose of this set of questions was to model a sequence of questions that might be asked of students as they read a narrative or, preferably, be asked by the students themselves. The passage is taken from one employed by Stephen Norris and Linda Phillips.[1]

Reading experts generally agree that reading involves more than decoding the author's meaning. Rather, readers construct their ideas as they progress through the text. This stress on construction of meaning implies activity. The reader must be aware of the kinds of information that she needs and to what extent her own knowledge is relevant to the text. The first question, which asked students to list some questions that would help them better understand the story, attempted to measure students' ability to recognize relevant kinds of information.

Part 2

The net was hard to pull. The heavy sea and strong tide made it even more difficult for the girdie. The good catch encouraged us to try harder.

What do you think is going on in this story? Why do you think so?

Major Concepts/Abilities Tested

Inferential comprehension

Correct Answers

*** 75% Correct inference based on context clues.**

"I think that they are fishing and they have caught something big and they really want to see what it is. I think this because of how it says, the net was hard to pull and the good catch encourages us to try harder."

Incorrect Answers

14% Correct inference, no reasons given.

10% Incorrect.

1% Blank.

Comments

Can students use contextual clues to construct a general meaning from a text, despite the use of unfamiliar words (e.g., girdie)? The difficulty of a text can be a function of vocabulary, but often there are other factors, such as the complexity of concepts or the structure itself, that account for whether or not students understand the meaning. This question measured students ability to use the context to overcome difficult vocabulary. Approximately three-quarters of the students were able to do this with justification. Those who gave no reasons for their inferences were generally poorer readers. Almost 30 percent of poorer readers (those at the 950 level of proficiency) gave no reasons for their inferences, another 17 percent either copied the story without explanation or continued questions elicited in previous question.

Part 3

Having reached our limit, we were now ready to leave. The skipper saw a threatening sky to the north.

The people in the story had "reached their limit." What does this mean?

Major Concepts/Abilities Tested

Inferring the meaning of phrases from context.

Correct Answers

- *** 12% They had caught all the fish they were allowed or the boat would hold.**
- *** 42% They had enough fish or they were tired of fishing.**
- *** 7% They had lost their patience.**
- *** 19% They had gone as far as they would go. (distance)**
- *** 4% Other reasonable inference.**

  "In the story, 'reached their limit' means their time is up. So they have to leave."

Incorrect Answers

- **13% Unreasonable inference.**

  "They are probably in a airplane and saw a storm to the north. Reached their limit means they are going as high as they can go."

- **3% Blank.**

Comments

Although a hallmark of good expository writing is its clarity, this is not true of narrative. The meaning that is given to ambiguous phrases can have important consequences to the general interpretation of the text. This question examined how students interpreted one such phrase.

The better readers were more likely to state that "the limit" referred to a legal limit or a physical limit (i.e., the capacity of the boat or a distance). A possible, although less plausible response, was that the fisherman had lost patience. Other interpretations were less convincing, suggesting that the students did not understand the meaning of the phrase.

What did the skipper see in the north? (Use your own words.)

Major Concepts/Abilities Tested

Understanding the meaning of a word from context?

Correct Answers

- *** 66% Interprets threatening as stormy, rainy, dark, etc.**

Incorrect Answers

- **12% Interprets threatening as mean, bad, etc.**
- **9% Repeats word "threatening."**
- **11% Another incorrect response.**
- **2% Blank.**

Comments

Although there were several plausible interpretations of the phrase "reached their limit," "threatening" holds a distinct meaning when it describes the sky, as opposed to a person. While 87 percent of above average readers were able to recognize this distinction, only 64 percent of average readers did so; while about 40 percent of poorer readers made no attempt to define the meaning of the word.

Part 4

We tied up to the wharf. We hastily grabbed our prongs and set to work. The sorting was done by the skipper. The boys did the cutting and gutting.

Why are the skipper and the boys doing what they are doing in this part?

Major Concepts/Abilities Tested

Inferential comprehension

Correct Answers

** 33% **Cleaning or preparing the fish.**

> "The skipper and the boys are doing post-fishing activities. The skipper is sorting types of fish catch and the boys are cutting and gutting. That is cutting the fish and taking out the parts you don't eat."

* 15% **Doing something because of the approaching storm.**

> "A storm was coming their way and they had to get out of there as fast as they could."

* 5% **Preparing the boat.**

> "The skipper and the boys are trying to fix their boat after a rough day at sea. They were also probably cleaning it with hoses."

Incorrect Answers

11% **"It's their job."**

> "Because it is a task they have to do, and so they don't have to do it later."

11% **Answer repeats wording from the passage.**

22% **Other incorrect.**

> "They were trying to build a lifeboat that they were going to row in away from that tornado."

4% **Blank.**

Comments

This question followed from the activities in the previous episodes. It is interesting to note that approximately one-quarter of the poorer readers answered "It's their job." This response suggests that these students did not understand that the question was asked within the context of the narrative itself and followed from the previous questions.

---

[1] Norris, Stephan and Linda Phillips. (1987) "Explanations of Reading Comprehension: Schema Theory and Critical Thinking Theory." *Teachers College Record.* Vol. 89, 2, 281-306.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 4

One Day When We Went Walking

One day when we went walking,
I found a dragon's tooth,
A dreadful dragon's tooth.
"A locust thorn," said Ruth.

One day when we went walking,
I found a brownie's shoe,
A brownie's button shoe.
"A dry pea pod," said Sue.

One day when we went walking,
I found a mermaid's fan,
A merry mermaid's fan.
"A scallop shell," said Dan.

One day when we went walking,
I found a fairy's dress,
A fairy's flannel dress.
"A mullein leaf," said Bess.

Next time that I go walking—
Unless I meet an elf,
A funny, friendly elf—
I'm going by myself!

Major Concepts/Abilities Tested

Inferential comprehension

The percentage of student responses and appropriate examples are listed under each of the categories below.

When the poet found a mermaid's fan, why did Dan say, "A scallop shell"?

Correct Answers

** 17% **Response indicates an understanding of the underlying theme of the poem by making explicit reference to Dan's lack of imagination in contrast to that of the speaker (poet).**

"The poet has a good imagination, but Dan is looking for a more real explanation. Dan is more of a person who is the type that thinks that there is a reasonable explanation for everything."

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

"The poet was creative, and liked thinking that one thing was like something else. But Dan didn't have a creative mind. He just thinks of things just like they are."

"Because the poet wanted to make the walk more imaginative."

*** 32% Response concentrates on the objects themselves but does not connect this to the larger issue of the difference between the viewers.**

"Dan said a scallop shell because even though it looks like a mermaid's fan it isn't. It's just a scallop shell and mermaids aren't real."

"Dan said, 'a scallop shell' because it wasn't a mermaid's fan, it was a scallop shell."

** *Full and complete response*
* *Adequate response*

Incorrect Answers

41% Response indicates a very literal interpretation.

"Because a scallop shell is shaped like a mermaid's fan and it also looks like any other fan."

"Dan said, 'A scallop shell' because that was what it looked like."

"Dan said a scallop shell because a mermaids fan would look like a scallop shell."

8% Other incorrect.

"Dan said, 'a scallop shell' because a mermaid probably uses a scallop shell for a fan."

1% Blank

Comments

The power of this poem depends upon understanding the contrast between two visions of the world: the fanciful, associative style of the poet as opposed to the realistic, objective style of the other characters. The first three stanzas set the scene, using the same structure, varying only in the details. In the last stanza the tense changes from past to future, and the mood changes from wistful to emphatic. However, without realizing the contrast in vision that is built up in the first stanzas, the point of the poem makes little sense.

Only a relatively small proportion of the students were able to express the general theme of the poem, i.e., the contrast between the imaginative and realistic perceptions of the characters. The majority of correct responses answered in more concrete and specific terms, referring to the objects themselves. Half the students either did not understand the poem or were unable to express their understanding.

If Ruth and Sue were walking together, what do you think their walk would be like?

Major Concepts/Abilities Tested

Ability to predict actions or reactions of the characters

Correct Answers

**** 13% Response describes the more prosaic approach of Ruth and Sue and gives a reason why.**

"Boring, because they don't have a good imagination."

"They would probably get along very well. Sue doesn't have a very good imagination, neither does Ruth, so the walk would be very real, and it wouldn't be as exciting."

"They wouldn't have a good time because they don't have much of an imagination."

*** 13% Answer as above but doesn't explain why.**

"If Ruth and Sue went walking together their walk would probably be more realistic than their walk with the poet."

"If Ruth and Sue went walking together the whole walk would be true."

Incorrect Answers

8% "Argumentative," etc. based on the responses of Ruth and Sue to the poet (speaker).

"They would have a hard time because they would not agree with nothing."

"They might find different things and maybe argue over what it was and agree."

8% "Interesting," based on the items mentioned in the poem.

"Great because when they go walking they'll find lots of things."

"It would be like a treasure hunt because they would find weird things."

8% "Interesting," based on a misconception of their personalities.

"I think if Ruth and Sue went walking together their walks would be very imaginative."

20% "Interesting," "funny," "scary," with no support.

"It would be a very strange and spooky walk!"

"I think it would be funny because they'd talk about what they thought things were."

8% Answer misses point and refers to details in stanzas 1 and 2 as support.

"It would be a very weird one because one has a dragons tooth and the other has a brownies shoe."

"If Ruth and Sue went walking, Ruth would look for things that look like dragons body parts. And Sue would look for a brownies shoe."

19% Other incorrect.

"They wouldn't have found as many things."

"Their walk would be like a road to a fairy tail."

2% Blank

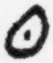
Comment

This question again investigated students' ability to understand the underlying theme of the poem by asking the students for a prediction. Most students found this to be a more difficult task. Not only were they required to understand the major theme of the poem, they had to interpret the characters of Sue and Ruth, and to project how they might interact with one another. Their responses clearly showed that students did not recognize the author's overall intent, but were distracted by the wealth of objects that were contrasted in the poem.

ON THEIR OWN

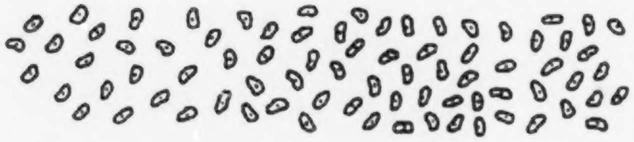
Students' Responses to Open-Ended Questions in Math

Grade 4

This is an oog: 

A package of oogs holds six of them. How many packages are needed to hold all the oogs pictured below? (HINT: Use the picture.)

The number of packages needed is _____.



Major Concepts/Abilities Tested

Numeration: concept of division

Ability to model the process of division

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answer/Method

**	50%	**13: with evidence that student partitioned into groups.**
**	12%	**13: with evidence that student counted and divided.**
*	9%	**Incorrect answer: evidence that student partitioned.**

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

* 7% **Incorrect answer: evidence that student counted and divided.**

8% **13: no work shown.**

** *good response*
* *adequate response*

Incorrect Answer/Method

14% **Incorrect answer: no method shown.**

1% **Blank**

Comments

Seventy percent of the students calculated that 13 packages were needed for the oogs drawn in the picture. Although not directed to do so, the majority (59 percent) spontaneously drew circles around groups of six and counted the groups. Most were successful using this method. Those who counted and divided (19 percent) and those who showed no work (21 percent) had more difficulty in calculating the correct answer.

The number of oogs was deliberately large in order to encourage children to order the data in some way. For those who understood the concept of division as a method of grouping, the task was relatively easy. For those who saw it purely in terms of the computation of numbers, it was more tedious and more vulnerable to counting and calculation error.

As with the story problem, this kind of task can be used for both instruction and evaluation. It focuses on the conceptual understanding of computation rather than the algorithm itself and shows a practical application of division as a way of simplifying data.

Find all the ways you can to fill in __ x __ = 36. List all your ways.

Now draw a picture, sketch, or chart to show one of your ways.

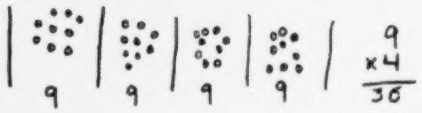
Major Concepts/Abilities Tested

Numeration: recognition of the factors of 36

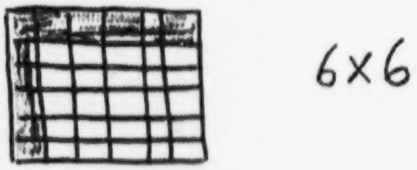
Ability to represent a computation in graphical form

Correct Answers

** 57% **Drew accurate and correct representation of multiplication.**

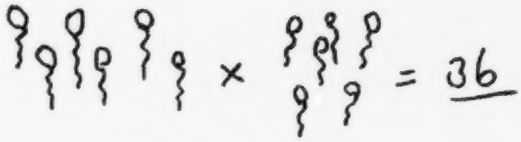


** 6% **Drew table, not necessarily filled in.**

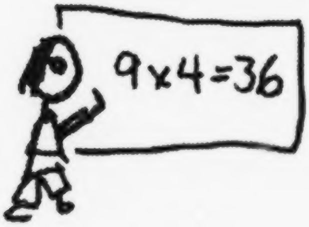


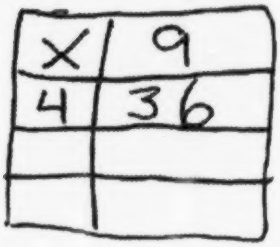
Incorrect Answers

13% Draws a representation of the number fact rather than the operation of multiplication.



20% Other incorrect answers.





X	9
4	36

4% Blank

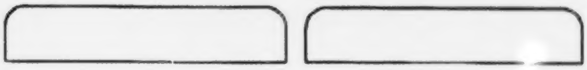
Comments

This question is essentially the inverse of the first task. Both require an understanding that multiplication and division represent operations performed on groups or collections. Although the two questions appeared in different forms, it should not be surprising that a similar percentage of students answered correctly in both instances.

Forty-two percent of the students could supply only three multiplication number sentences (3 X 12; 4 X 9; and 6 X 6); while 12 percent gave all nine possible combination of factors. Forty-three percent used the commutative property in listing their combinations. The hurdle for most children was 18 X 2 and 36 X 1 as possibilities.

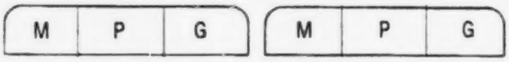
In response to the second question, approximately 60 percent were able to represent multiplication as a number of sets. The remainder of the children did not appear to understand how multiplication could be represented in any concrete form, despite their ability to supply the correct multiplication facts.

Mary, Paula, and George want to share 2 candy bars equally. Show in the drawing below how much candy each person gets. You may use the initials of the children (M, P, G) if you wish.

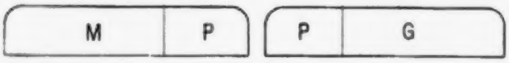


Correct Answers/Methods

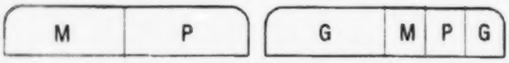
** 45% Shows each candy bar divided into three, equal pieces.



* 12% Shows each bar divided into two pieces—1/3 and 2/3.



* 3% Other correct answer.



Incorrect Answers

5%	Divides each bar in half.
4%	Candy bars unequally divided.
8%	Other incorrect response
22%	Blank

Now, write a fraction for the portion of candy each child gets: ____.

Correct Answers

**	27%	1/3
*	18%	2/6

Incorrect Answers

18%	2/3
9%	1/2 or equivalent
25%	Other incorrect
3%	Blank

Comments

This question repeats the theme of the other two—the representation of number operations—in the context of fractions. Despite the greater complexity of the task, the proportion of children who were able to draw an accurate representation remains similar to that found in the whole number problems. Not surprisingly, fewer were able to name the fraction that they had just represented. However, it appears that a necessary condition for success in writing the fraction is the ability to visualize how the candy bar would be divided. Children who could not do this were at loss to understand the computation process involved.

As a whole, the three questions illustrate the importance of representation as a way of understanding what operations actually mean. In the introductory stage of a topic, such as the case of fractions, they provide a context for understanding how "they work." In the case of familiar topics, such as whole number operations, questions such as these are easy ways to evaluate how well the student actually understands what is being done. Once students become proficient with their "number facts" it is easy to assume that they understand what these facts represent. Questions such as the ones given above are easy ways of checking such assumptions.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 4

The percentage of student responses and appropriate examples are listed under each of the categories below.

Write a story problem that could be solved by the computation below.

$$\begin{array}{r} 12 \\ \times 4 \\ \hline 48 \end{array}$$

Major Concepts/Abilities Tested

Numeration: concept of multiplication

Ability to recognize the function of the multiplication algorithm in an everyday context

Other Concepts/Abilities Tested

Communication of mathematical ideas

Understanding of mathematical language and symbols

Correct Answers

** 52% **Appropriate word problem suggesting multiplication.**

"Tommy is going to buy four rugs. Each rug costs $12.00. How much money did he spend in all?"

** 2% **Appropriate word problem suggesting repeated addition.**

"Meghan has 12 stamps. Jon gives Meghan 12 more stamps. Then Karen and Tom each given Meghan 12 stamps apiece. Meghan wants to know how many stamps she has. What should she do?"

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

** 3% **Gives problem the answer of which is a factor, not a product.**

"I had 48 apples. I divided them equally among my 12 friends. How many did they each get? They each got 4 apples."

*9% **Weak problem situation.**

"Tara took 4 cookies, 12 times in a row on Monday. How many cookies did Tara take in all?"

** *good answer*
* *acceptable answer*

Incorrect Answer

7% **Explains procedure for multiplying 4 X 12.**

"First we take the two from the twelve and times it eight and that = 8. And then we take the 1 and times 4 and that = 4. And then you put it in front of the eight and that = 48."

"The problem is 12 times 4. You could add 12 4 times. The answer is 48."

5% **Problem is vague or unclear.**

"In 2 days there are 48 hours. In one day there are 24 hours. In one half of a day, there are 12 hours. You know that in two days there are 48 hours. You know that because 12 X 4 = 48."

4% **Gives addition problem.**

"12 students went to their class. 4 more kids went to their class late. How many kids went to class? 12 X 4= 48."

"Josh had 12 stamps. Mike gave Josh 4 stamps. How many did Josh have now. Hint—multiplication."

2% **Gives problem in which there are "4 times as many."**

"On Sunday afternoon there were going to be football try-outs at 12:00 p.m. At 11:45 a.m. 12 kids showed up, but between 11:45 a.m. and 12:50 a.m. 4 times as many kids showed up than at 11:45 a.m. How many kids showed up in all?"

20% **Other.**

"Carrie had 12 baseball cards. Maudie had 36. How many baseball cards in all?"

"A man put 12 cans of tomatoes on the shelf. By the end of the day there were only 4 cans of tomatoes left on the shelf. How many cans were on the shelf in the beginning?"

2% **Blank**

Comments

Over half the students were able to give an appropriate and convincing problem for the computation given. The other students were less sure about the kinds of situations in which multiplication would be useful. Some students (as illustrated above) appeared to confuse multiplication with addition. Others used phrases such as "four times as many" without thinking about the language itself. Although it is true that 48 is "four times as many" as 12, this phrase does not signify the same thing in a problem.

Write one story problem that takes 2 steps to solve. The steps must use the computations shown.

$$\begin{array}{r} 35 \\ \times\, 2 \\ \hline 70 \end{array} \qquad \begin{array}{r} 70 \\ +\, 25 \\ \hline 95 \end{array}$$

Correct Answer

** 35% **Appropriate multi-step problem.**

"Each of the king's light infantry had 2 weapons, a handaxe, and a short sword. The king has 35 light infantry. How many weapons do all the light infantry have together?

Each of the king's 25 heavy infantry have 1 weapon, a battle axe. The king's army has only heavy and light infantry, no knights or archers or anything else. How many weapons does the king's army wield?"

"Jill had 35 stamps on each page. There were two pages. Then she bought 25 more stamps. How many stamps does Jill have?"

* 3% **Appropriate multi-step problem using different numbers from the ones given.**

"Diana has 20 oranges in one box. She has two boxes. She also has 15 oranges in another box. How many oranges does she have in all?"

* 5% **Appropriate multi-step problem, but omits question or asks inappropriate question.**

"Tim had a dollar. He bought 2 things for $.35. Then he bought something for $.25. Does he have enough money?"

"Robyn and Jeff had $35.00 each. David had $25.00 more than them put together."

** *good answer*
* *acceptable answer*

Incorrect Answer

10% **Writes computation in words or explains how to solve problems.**

"Well, 35 X 2 = 70. How? Well, 5 X 2 = 10, right? Yes. Bring the one up over the 3, and 3 X 2 = 7."

8% **Two separate, unrelated problems.**

"There's 35 people and 70 cookies. If each one has two cookies each, how could you solve it? If there's 70 kids and 25 more enter, there will be 95 kids in all."

6% **Problem situation vague or unclear.**

"I went to the store and bought 70 bags of chips and 95 sodas."

29% **Other incorrect.**

"Mark wanted to buy 2 apples. Both of them costs 35 cents. Then he wants to buy 25 pieces of gum. The gum costs 70 cents for all 25 pieces. (70 + 25 = 95. 95 + 70 = 165)"

"Jane had 70 cookies. Then she bought 25 more. Then she wanted more so she bought 2 more packs of 35 cookies. How many cookies does she have now? (35 X 2 = 70. 70 + 70 = 140)"

"One day Jill had 35 brouchers. Her mom gave her 2 more. Then she had 70. Then her mom gave her 25 more. How many did she have then?"

4% **Blank**

Comments

Not surprisingly, a two-step computation presented more difficulty than a simple multiplication. However, the difference in the number of students who succeeded in translating the calculations into everyday problems was not great enough to suggest that the difficulty with such questions lay in the computation itself. It appears more likely that the students who could not answer either question correctly (approximately 40 percent) do not understand the meaning of the calculations that they perform.

The difficulties that students experience with translating word problems into arithmetic procedures have been well documented and are often laid to difficulties in reading comprehension. The results above suggest that different factors may be involved. If children cannot place their arithmetic calculations into some kind of context that is meaningful to them, it should be no surprise that they find difficulty in trying to decipher what the author of the problem intends. In other words, if reading is a process of constructing meaning from the text, as suggested by many educators, it is as important for children to understand the arithmetic concepts as it is for them to understand the verbal ones.

Problems such as the ones above place control in the hands of the students. It gives them the opportunity to think of the effect of operations on numbers and how these operations can be used to order events in their own lives. Asking students to think about numbers in context also helps them to understand the algorithms involved. How does "six more" differ from "six times as many"? The language of mathematics can be difficult for students at this age. By relating mathematical language to everyday events, students become more facile in communicating their ideas mathematically. Finally, this kind of problem offers a good opportunity for group activity as students create and critique possible scenarios in which numerical operations play a part.

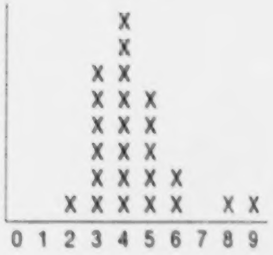
ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

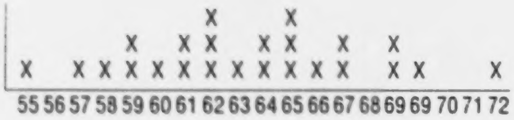
Grade 4

*Here are five graphs. They show different things about a class of fourth graders and their families. Look carefully at these graphs to answer each of the three questions below.**

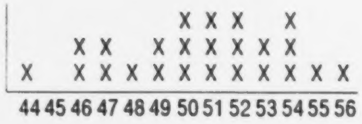
GRAPH A



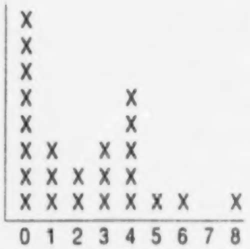
GRAPH B



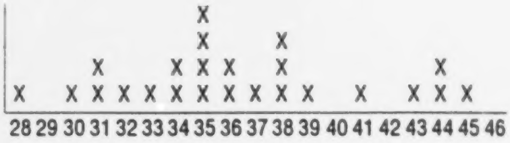
GRAPH C



GRAPH D



GRAPH E



---

* Adapted from a set of questions developed by TERC (Technical Education Research Center).

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

Which graph shows the number of cavities that the fourth graders have? ____ Explain why you think so.

Major Concepts/Abilities Tested

Interpretation of graphical data

Reasoning and communication

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers/Methods

*** 19% **Selects D and gives strong supporting argument.**

"Fourth graders wouldn't have that many cavities so therefore it would be a graph with lower numbers. So it couldn't be graph B, D or E. Graph A and D are each small number graphs, but Graph A is a little too large. So it's Graph D."

"Graph D. I think this because most kids my age don't have many because there are new things that keep us from getting them. I know it is not graph A. I don't have any and most of my friends don't."

** 30% **Selects either D or A and gives argument with some merit.**

"Graph D. It has low numbers."

* 26% **Correct graph, but poor or no argument.**

"Graph D. Because each X could stand for one kid who had that many cavities."

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers/Methods

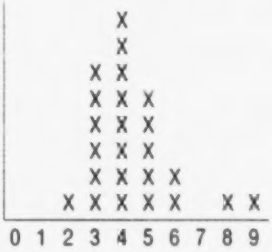
20% **Other incorrect response.**

"Graph B. I think it is B because a lot of kids eat candy."

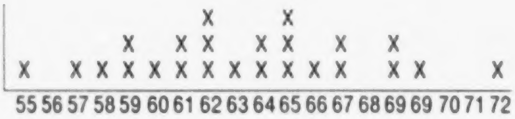
"Graph C. I think this because in graph A, there are too many X's to be cavities. In graph B, the graph is too long. In graph D, there are too many X's, and graph E has the same reason."

3% **Blank**

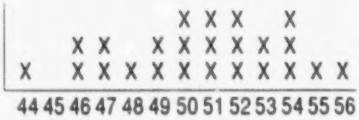
GRAPH A



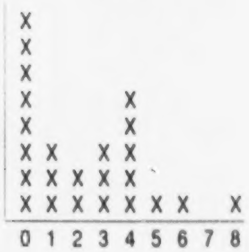
GRAPH B



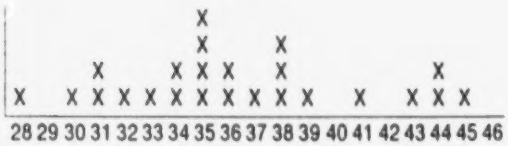
GRAPH C



GRAPH D



GRAPH E



Which graph shows the number of people in the fourth graders' families? ___ Explain why you think so.

Correct Answers/Methods

***** 15% Selects A with strong supporting argument.**

"Graph A. It's impossible to have a family of 0. For fourth graders it's impossible to have a family of 1. Also 4 is a balanced number and the most of fourth grade families is at 4."

"Graph A. I chose this because of the following reasons: no one that has a family has 0 family members, no one (except an orphan) has 1 family member, 2 is common for divorced or separated parents, also not many people have 8 family members."

**** 28% Selects A or D and gives argument with some merit.**

"Graph A. Families are usually made up of 2 to approximately 8 people. The highest number of x's is 8. So my guess is as affore stated."

*** 2% Explanation suggests extended families.**

"Graph B. Because everybody has to have cousins, aunts, and uncles."

25% Correct graph, but poor or no argument.

"Graph A. Because it has the highest amount of families on the graph."

Incorrect Answers/Methods

24% Incorrect response.

"Graph D. I think the highest number of people in a fourth grade family is 4."

"Graph D because graph A shows the number of cavities in fourth grade."

4% Blank

Which graph shows the heights (in inches) of the fourth graders? ___ Explain why you think so.

Correct Answers/Methods

***** 14% Selects C and gives strong supporting argument.**

"Graph C. I am 48 inches tall which is normal in 4th grade. And Graph C shows 48 inches on it, and no other graph shows that number which is 48."

**** 24% Selects either C, B or E and gives argument with some merit.**

"Graph C. No one would be seventy two inches high. Fifty six is right for most fourth graders."

"Graph C because they're all around the same height."

*** 31% Selects either C, B or E, poor or no argument.**

"Graph B because most fourth graders are tall."

"Graph B. Since we eat better food today I think B would be correct because we are getting much taller."

Incorrect Answers/Methods

23% Other incorrect response.

"Graph E shows different sizes and fourth graders are different sizes."

5% Blank

Comments

The greatest amount of credit was given to those students who indicated that they had read and interpreted the data on each graph before coming to a conclusion. In practice, that meant that they made their argument in terms of evaluation and comparison. Those who received two marks indicated in their response that they had read and understood the graph which they chose. Although their choice was not necessarily correct, it was reasonable. One credit was awarded students who chose the correct graph, but gave a minimal response. It was assumed that they had gone through some process of evaluation, but were unable to verbalize their reasoning.

Undoubtedly, this type of question is unfamiliar to students, and their results reflect this. Less than 20 percent were able to defend their choice of any graph with a strong argument. However, the questions do not represent material that is outside the fourth grade curriculum. Fourth graders often make their own frequency charts of things or events. They are also expected to read graphs and to answer questions about them. What makes this type of question different is the requirement that they apply their general knowledge to a mathematical situation and to ask, "What is reasonable?" Not only is this a productive activity for class or group discussion, but it reinforces the notion that numbers do not exist in a vacuum but represent situations that occur in the actual, everyday world.

ON THEIR OWN

Students' Responses to Open-Ended Questions
In
Math, Reading, Science, Social Studies

Results of the 1990 Assessment

Grade Eight

Massachusetts Department of Education
Massachusetts Educational Assessment Program

BEST COPY AVAILABLE

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Science Background

Currently, science curriculum is undergoing serious examination in light of recent findings indicating that United States students scored near the bottom in various international science exams. In response, studies (Science for All Americans, Research Within Reach: Science Education; 1994 NAEP Science Consensus) have generated curricular recommendations that are remarkably similar. Among those recommendations are:

Coverage should be in-depth.

By covering one topic thoroughly, students pick up skills that can be applied to other topics; they develop the habit of science. By stressing one concept and reinforcing that concept, students gain a depth of understanding than is far more substantial than a group of disparate facts. In other words, in the teaching of science, less is more.

Science should be taught as cross-discipline and multi-discipline.

Science is not a single area, but a series of related topics. Many of the concepts cross areas. The different subject areas do not exist in a vacuum and the interrelationship of the various disciplines provide a totality that can be offered in no other manner. Particularly at the secondary level, science is taught as series of discrete disciplines. Nor is science isolated from other disciplines. The relationship between science and mathematics is obvious, but science and the

humanities, the fine arts, and the social sciences are also intrinsically tied to one another.

Science education should begin early.

Elementary teachers might be tempted to give short shrift to science because they may feel they don't have the training to teach science, yet early science education is vital to developing students who have a love and respect for science. We know that despite much instruction, students often retain misconceptions particularly about the natural world. The earlier teachers can begin to delve into these misconceptions, the earlier they can start correcting them.

Science should begin with questions, not with answers.

Instead of teaching facts to students, a more engaging way of teaching might be to ask "Why do you think...?" A major characteristic of scientists is curiosity. This approach nurtures a sense of curiosity and the resulting sense of accomplishment that accompanies it. Answers should not be definitive but rather lead into more and deeper questions.

Science should be done, not read or heard.

Science is a series of discoveries based on careful observation and calibration. How better to actively engage students in science than with a hands-on approach, particularly in the early years. Students who begin with hands-on science develop facilities for observation, handling equipment and developing hypotheses that can be transferred to other subject areas. Doing science is more than following a pre-selected experiment with orderly steps leading to a predetermined conclusion; it is observing phenomena, developing a hypothesis, testing that hypothesis, and replicating that test in order to confirm the results.

Science should be cooperative.

Since science in the real world is a cooperative activity, it makes sense that science in school should be taught in the same manner. A scientist does not work in isolation, but shares goals and data. The scientific community tends to share knowledge and resources for the greater good of the scientific world, hence the world at large.

Science should have a practical application.

Unless students can assign a personal value to science, it remains arcane and distant. Using science helps students to appreciate its relevance in their own lives. Knowing science can only help consumers and voters make educated decisions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, students were asked to explain why or how something happened. In the area of earth science, they were asked to explain how water forms a canyon; in physical science, they were asked to explain condensation and evaporation. Among higher order thinking skills, students were asked to come up with a model of a natural phenomenon. They also had to complete the design of an experiment.

Teachers were given a selection of the open-ended questions prior to the test and asked to evaluate how well prepared their students were to answer the questions. At grade four, the relationship between actual performance on the questions and teachers' judgements of student preparedness for the questions varied. One-third of the teachers felt that their students were very well prepared to answer a question about evaporation and condensation, while eight percent of the students were able to discuss both evaporation and condensation successfully. On a question that asked students to consider characteristics of birds in relation to their habitats or food, 32% of the teachers felt that their students were very well prepared to answer. However, less than 15% of the students gave answers that were considered complete.

Grade Eight

The grade eight questions ranged from physics (laws of motion and friction and gravity) to physical science (the effects of temperature on a substance) to earth science (the causes of earthquakes) to scientific inquiry and process skills (steps in separating sand and sugar; reducing data in order to generalize;

generating an experiment). Other concepts included adaptation and the nature of sound waves.

As at grade four, teachers were given a small selection of questions and asked to evaluate their students' preparedness in responding to the task. Forty percent of the teachers felt that their students were very well prepared to describe an experiment, including the statement of hypothesis, design and necessary data. Approximately 15% of the students answered this question successfully. The largest discrepancy between teacher predictions and student performance was in response to the question "Explain why the moon does not fall to earth." Almost one-third of the teachers (31%) felt that their students were well prepared to answer this question, yet only 4% of the students could answer it successfully, with another 12% able to answer it with partial success.

Grade Twelve

At grade twelve, the questions focussed on interpreting data and making predictions including graphic representations. Other questions involved asking survey questions and generating a hypothesis. Students were also asked to explain why the moon doesn't fall to earth.

As in the case of the other grades, 12th grade science teachers were given the opportunity to consider the questions and respond to their students' preparedness. Given a scenario of fish dying unexpectedly, over one-fourth of the teachers predicted that their students were well prepared to generate two hypotheses and appropriate tests. About one-fourth of the students were able to generate good hypotheses and test them accordingly. Over one-third of teachers (35%) stated that their students were very well prepared to explain why the moon does not fall to earth, while 11% answered successfully. This question was problematic at both grades, with the teachers predicting much greater student preparedness than was the case. It may be that teachers felt they had covered the subject sufficiently.

At both grades eight and twelve, the majority of teachers (63% at grade 12 and 59% at grade 8) listed student indifference and lack of motivation as a very important reason why students do not perform well in science. And over 40% at both grades listed lack of preparation prior to entering the class as a cause of low achievement in science.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Math Background

In 1990 the National Council for Teachers of Mathematics published ***The Curriculum and Evaluation Standards***, which challenged the traditional building block model of mathematics learning. According to this model, school mathematics consists of a set of discrete skills or building blocks—computational algorithms, measurement formulae, and algebraic rules. The elementary school teacher sets the first layer of blocks, the middle school teacher the second layer, and so on until, eventually the edifice is large and strong enough for students to be said to know mathematics. Although problem solving is a goal of learning, it is not actually connected to instruction. It comes later as a natural consequence of learning. The role of the teacher is to ensure that students solidify their learning at each level in order to progress to the next until, ultimately, they would be able to "do mathematics," i.e., solve real problems.

Unfortunately, for most students, this model is not working. Students are not automatically solving problems. They are not making the connections. Furthermore, most are getting stuck at the lower levels, so they are not even allowed the chance to meet real problems.

In contrast, the Standards suggested a different model for mathematics learning—one of a network of skills and concepts that are related to one another. Students can go in and out of the network at any level because the ideas are connected. Furthermore, problems form

the context of this network, because solving problems is how we begin to understand what mathematics is about.

In their publication of the ***Standards***, as well as the more recent ***Professional Standards for Teaching Mathematics***, the National Council for Teachers of Mathematics stresses understanding over memory, relevance over training. It advocates an expansion of the curriculum to include new topics—such as probability and statistics, patterns and relationships, discrete mathematics—that reflect the mathematical needs of society. It gives greater emphasis to other topics—such as geometry and measurement—that have often been relegated to the back of the book. It delineates four basic abilities that are essential to all topics: reasoning, communication, problem solving, and the ability to see the connections between mathematical concepts. By including evaluation, it makes the point that assessment should be viewed as an integral part of instruction.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

The open-ended questions were designed to cover different important concepts and abilities in the fourth grade mathematics curriculum. Questions concentrated on numerical operations and the ability to extract and use information presented in graphical form. There was little stress on calculation itself. Instead, students were asked to represent multiplication and division by writing stories and drawing pictures. They were required to interpret the data that were presented in a set of graphs and were asked to organize data. A final set of questions dealt with spatial visualization and geometric shapes.

When given a small selection of the actual questions, less than a quarter of the teachers replied that they believed their students to be well prepared to answer most; although another half believed that they were somewhat prepared. The exception was in response to a question asking students to write a story problem for a simple computation. In this case, 40 percent of the teachers replied that their students were well prepared.

In general, fourth graders were not hesitant to answer questions that may have been unfamiliar to them. In most cases, their answers showed a willingness to apply their reasoning and general knowledge to the task.

Grade Eight

The set of questions at this level contained both familiar word problems involving percentages and fractions and less common graphical problems that required students to extract and synthesize different types of information. As in the case of the fourth grade, teachers were presented with a selection of the problems and asked to respond in terms of their students preparation. It is not surprising that many teachers (between 45% and 71%) believed that their students were very well prepared to answer the word problems and those that probed students' understanding of numerical calculations. In contrast, only 20 percent responded that their students were prepared to use a map and a time/distance graph to describe a journey. Other questions—one concerning number patterns and one that asked students to synthesize and apply information from two graphs—were positively responded to by approximately a quarter of the teachers.

At this level, students' achievement reflected teaching more than in the earlier grade. Although few questions were answered well by more than 40 percent of the students, those tended to be the more familiar computational word problems. In response to these questions, high achievers on the multiple-choice section of the test were more successful than others. However, all students performed poorly on questions that asked them to observe, reason, and play with numbers. For example, there were few students who were able to draw inferences from graphical information or to identify more than one number pattern. In general, the eighth grade students were less prepared than their younger brothers and sisters to apply their reasoning in a mathematical context that was unfamiliar to them.

Grade Twelve

Again, at grade twelve, the preponderance of questions were in the area of numeration, with some questions focused on statistics and one which measured students' understanding of area. In most cases students were asked to explain or justify their response. Answers were judged on the quality of their response and credit was given for method in addition to the solution itself.

As was the case with the other grades, teachers were presented with a selection of questions and asked to comment in terms of their students' preparation. However, teachers who answered the twelfth grade questionnaire replied with reference to all grades at the high school level. With the sole exception of the value of 5!, no more than 20 percent felt that their students had been well prepared to answer any of this sample of questions. It seems that the written justification and reasoning required by these questions is not part of the usual high school curriculum.

Summary

From the teachers' point of view, students are not well prepared to answer questions that require them to synthesize information from various sources, to discover patterns, or to draw conclusions from data and argue their point of view. In general, the younger children—whether prepared or not—were more disposed to do so, and many performed well. As students progress through the grades, they appeared to be correspondingly less willing to attempt to answer questions that go beyond the specifics of the taught curriculum or required them to apply or discuss the concepts involved.

At grades 8 and 12 familiar questions that reflected the curriculum (e.g., the words problems at grade 8; questions that involved algebra at grade 12) were correlated with performance on the multiple-choice section of the test. The relationship between multiple-choice achievement and achievement with the less conventional questions was weaker, partially because few students at all proficiency levels performed well, partially because the explanations of the best students was not consistently better than those less capable.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.

- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.

- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Social Studies Background

Currently, social studies is experiencing a "back to basics" movement stressing history, civics, and geography. This trend portends new curriculum examination and revitalization. Recent research has indicated that in elementary and middle grades, teaches make false assumptions about what their students know. Textbooks contribute to this problem by making the same assumption. In a recent article by Margaret McKeown and Isabel Beck,[1] researchers considered the background knowledge that middle school students bring to the study of history, the American Revolution in particular. They say that

> "elementary social studies textbooks seem to assume knowledge of more specific concepts than students have under control...students do seem to have a general landscape involving the country being settled, the seeking of freedom and some notions of a process...Yet this seems exactly where the texts focus their material."

This miscalculation of student knowledge calls for more complete and in-depth assessments of what significant knowledge is lacking. The longer that misconceptions continue uncorrected, the more difficult it becomes to correct them. Students will integrate knowledge to accommodate misinformation.

According to McKeown and Beck, one of the greatest challenges facing educators is finding and correcting

the confusions that young students bring to class. They suggest using a framework of the topic to organize previous knowledge and new information into a coherent whole. By presenting the relevant information in a framework, students would be better able to see the totality of the topic and evaluate whether their previous knowledge will fit in. They might be able to recognize irrelevant or inaccurate information using the framework as a sounding board.

Retention of knowledge is a problem at the secondary level. Social studies must build on the knowledge structures of the past. All this calls for a unified curriculum that considers how and when to teach materials. Walter Parker in *Renewing the Social Studies Curriculum*[2] suggests that themes in social studies can begin at the elementary grades and be reinforced across the years; that is, "content at any grade level should be presented in ways that provide, insofar as possible, a comprehensive view of a complex whole." This suggests a thematic approach without sacrificing chronology or coverage, which would satisfy those who call for social studies to have breadth over depth.

Parker also discusses attributes of a successful social studies classroom.

- Lessons have substantive coherence and continuity. Disparate facts are pulled together.
- Students are given time to think—to prepare responses to questions.
- The teacher asks challenging questions and structures challenging tasks.
- Students are able to offer explanations and reasons for their conclusions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, the open-ended questions covered process skills such as map reading and interpreting graphic representations. Another question focussed on a sense of history. Other questions involved conservation and economics.

Fourth grade teachers were given a small selection of these questions and asked to predict how prepared their students were to answer them successfully. Over half the teachers felt that their students were very well prepared to compare their own lives with those of colonial school children, while 33% of the students were able to do so. In response to how well their students were prepared to write a set of directions using a map, 31% of the teachers felt that their students were very well prepared. One-quarter of the students were able to do so.

Grade Eight

Eighth grade students were asked to interpret a political cartoon. A series of questions asked students to evaluate the quality of several arguments. Both specifics of history (students were asked to place six events on a time line) and a sense of history (students had to evaluate evidence to identify a speaker and when he lived) were covered. Students were asked to generate three courses of action and their negative impacts in response to an environmental crisis.

As in grade four, social studies teachers were given a small selection of open-ended questions and asked to evaluate how prepared their students were to answer those questions. About 15% of the teachers felt their students were very well prepared to generate three courses of action and their corresponding negative impacts; approximately 20% of the students were able to do so. When asked how well prepared their students were to explain why an argument was strong, about half the teachers felt that their students were very well prepared to do so. Eleven percent of the students were able to interpret and explain the strength of an argument.

Grade Twelve

Seniors were asked to interpret population graphs and make predictions based on those graphs. They were also asked to discuss how the policy of open immigration has affected the United States socially, politically, and economically. Another series of questions asked students to interpret two different graphs of immigration rates and discuss changes the graphs illustrated. As in grade eight, seniors were asked to interpret a political cartoon, this one about gun control.

Teachers who taught twelfth grade social studies were asked to comment on the preparedness of their students when given a sample of the questions. About one fourth (27%) of the teachers felt that their students were very well prepared to discuss the political, economic, and social effects of immigration while no more than 13% of the students were able to give complete answers in any of the categories. When given two graphs of immigration rates, one spanning 155 years, the other one year, 82% of the teachers felt their students would be able to answer why the numbers were larger in one graph than the other. Thirty-five percent of the seniors were able to answer that one graph spanned a much longer period of time than the other.

At grades eight and twelve, it appears that teachers felt their students were better prepared than they were to answer those questions that dealt with higher order thinking skills, such as evaluating arguments (grade eight) and cause and effect (effects of immigration at grade twelve).

In grades eight and twelve, the majority of teachers cited student indifference and lack of motivation as important factors in students' lack of success in social studies. Almost half the teachers at both grades cited poor preparation prior to entering the class as another important factor in students' lack of success.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

[1] McKeown, Margaret G. And Beck, Isabel L. (Winter 1990). "The Assessment & Characterization of Young Learners' Knowledge of a Topic in History." American Educational Research Journal. Vol. 27, 4, pp. 688-726.

[2] Parker, Walter C. (1991). ***Reviewing the Social Studies Curriculum.*** Alexandria, Virginia: Association for Supervision and Curriculum Development.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Reading Background

During the past few years, the ideas that were expressed in ***Reading and Thinking*** have been refined and expanded by educators and philosophers interested in the reading process. Different themes have emerged from these activities.

One stresses the constantly shifting attitudes that the reader assumes in trying to understand any text. In contrast to the notion of reading as the accumulation of information, Judith Langer[1] proposes four different kinds of relationships that occur and reoccur during the process of reading.

Being out and stepping in. Readers use the information from the text and their background knowledge to get enough information to "step into" the author's vision. In literature, readers try to make initial acquaintance with the character, plot and setting; while in exposition they try to figure out what the topic is about.

Being in and moving through. Readers immerse themselves in the author's vision, trying to understand the meaning of the author. In exposition, readers take each new bit of information, try to understand it and to link it to what they already understand the text to have said about the topic. In fiction, readers use each new bit of information to go beyond what they already understand—asking questions about motivation, causality, and implications.

Being in and stepping out. Readers relate the text to their own knowledge and experiences. Readers of fiction use what they read in the text to reflect on their own lives, on the lives of others, or on the human condition in general. In non-fiction, readers use the text information to rethink information they already know.

Stepping out and going beyond. Readers distance themselves from the text and assume a critical stance, judging the text and relating it to other texts or experiences.

A second theme has suggested that literature is a powerful context for the teaching and learning of critical thought. As readers construct their understanding of the text, interpretations are often not possible. In fact, as Stephen Norris and Linda Phillips[2] suggest, the essence of critical reading is to raise alternative interpretations, weed out interpretations to the extent that available information will allow, and then remain with multiple possibilities. In their view, literary thinking is a complex reasoning process involving the analysis, synthesis, reformulation, linking, and generalization of ideas.

One important implication that can be drawn from this reappraisal of the reading process is that evaluation cannot be made purely on the basis of whether or not the reader's conclusions are similar to those of the teacher or test constructor. Instead, it is the quality of the reader's argument or justification that is of central importance.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

Drawing inferences proved extremely difficult for fourth graders. Given a poem, students were asked to identify the underlying theme and to draw inferences about the characters. Although the context of the poem (children walking through the woods and finding objects) appeared to be appropriate for this age, the meaning of the poem depended upon an appreciation of different perceptions that people have. In this case, the poem contrasted the imaginative, fanciful vision of the poet to the more prosaic and realistic vision of the other characters. Few of the children understood this. Among even the best readers (those at the 1550 and 1850 proficiency levels), only approximately 30 percent were able to state the underlying theme of the poem. When asked to use their interpretation to predict the action of the characters, the number dropped to 25 percent.

Students also experienced difficulty in interpreting the terms of an advertisement that was aimed at children of this age. Although 43 percent of fourth grade teachers believed that their students were well prepared to answer this type of question, this particular task was particularly challenging. Cluttered layout and deliberately ambiguous prose led the students astray. Despite the fact that the more proficient readers at this grade level showed a well-developed ability to handle details in multiple-choice questions, only approximately 30 percent were able to interpret the terms of this advertisement correctly. When asked to assume a critical stance by identifying specific examples of persuasive techniques, most students at all levels succeeded to some degree. However, only a few were able to characterize and describe the effects of such techniques on readers.

Grade Eight

The greatest discrepancy between teacher expectation and student performance occurred in response to a question which required students to identify and judge a set of assumption. Whereas less than 20 percent of teachers stated that their students were well prepared to answer this type of question, the overall success rate among students was more than twice that figure. On the other hand, performance did not follow the usual pattern of increased success rate with increased proficiency levels. Almost half of the most proficient students identified a valid assumption but, instead of judging the reasonableness of the assumption, they commented on the argument itself. Obviously, teachers did not view this question as one that fell within the area of reading instruction. They may have seen it as within the purview of social studies. Students also may have interpreted the question as one requiring argument rather than analysis. However, as discussed previously, the role of analytic and evaluative skills are not confined to a specific subject area, but are part of any type of critical reading.

Grade Twelve

On the whole, the passages on which open-ended questions were based were longer and more demanding at grade 12 than at the preceding levels; however, similar abilities were tested.

Given a passage from the James Dickey's classic *Deliverance*, students were asked to infer the character of Lewis from description and dialogue. Sixteen percent effectively identified several traits or provided an in-depth discussion of one trait, while another 21 percent gave a less comprehensive response. There was a clear distinction in the quality of responses among the different proficiency levels. Forty-two percent of students who performed at the 1850 level in the multiple-choice section of the test gave an effective and convincing response, in contrast to 24 percent of students at the 1550 level, and many fewer at the lower levels.

The ability to infer meaning from text was tested in a practical context through the use of an article on teen-age traffic deaths. The article compared a number of disparate studies on drinking and driving and concluded that more stringent laws against drunk driving may be less effective than media coverage. Students were asked to identify the main point of the article and to justify their response. This proved to be challenging for even the most proficient readers. Only 21 percent of 1850 level and 12 percent at the 1550 level could give a clear and convincing argument.

The majority of students did not make the causal connection that was suggested by the article. Many stopped at the headline or made the general statements about drinking and driving without understanding the requirements for justifying their response. Few used the statistics in the article.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.

- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.

- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of students' understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

---

[1] Langer, Judith. (1989) ***The Process of Understanding Literature***. Center for the Learning and Teaching of Literature, 1400 Washington Avenue, University at Albany, Albany, NY 1222.

[2] Norris, Stephen and Linda Phillips. (1987) "Explanations of reading comprehension: schema theory and critical thinking theory." ***Teacher College Record***. Vol. 89, 2, pp. 282-306.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 8

Roger inflated a balloon, tied the end closed and put it into the refrigerator. What happened to the balloon? Did it get larger, get smaller, or stay the same? Explain.

Major Concepts/Abilities Tested

Physical science

Temperature and states of matter

Particulate nature of matter

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 3% **Smaller, with explanation at molecular level including that the molecules slow down and move closer together as the temperature decreases.**

** 4% **Smaller, cold causes molecules to slow.**

> "It probably got smaller because the air is cold and the molecules would stop moving."

** 17% **Smaller, cold causes molecules to contract.**

> "It got smaller because the colder a gas gets, the more it contracts. The molecules of air inside the balloon drew closer to each other taking up less space."

* 13% **Smaller, rubber or other unnamed substance contracts when cold.**

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

"The balloon gets smaller if you put it in the refrigerator because in heat it gets larger and that means in the cold it gets smaller."

(This response states that things contract in the cold without discussing why they contract.)

* 5% **Smaller, air escaped balloon.**

"The balloon got smaller. Most things contract as they get cold. The rubber of the balloon would get smaller forcing the air out. They balloon may have a hole in the first place letting oxygen seep out making it easier for the balloon to contract."

(This rather logical response deals with the balloon as a whole rather than its constituent substances.)

Incorrect Answers

24% **Smaller, no explanation or incorrect explanation.**

"It got smaller because the evaporation in the refrigerator made it smaller and because there is no air or oxygen in it."

(Although the student is partially correct here insofar as the balloon did get smaller, she does not receive credit for the response since her explanation indicates no understanding of the effects of temperature on a substance.)

14% **Same.**

"It would stay the same. The temperature wouldn't change the balloon because it is already closed."

11% **Larger.**

"The balloon got larger. Cold air expands and pushes the molecules farther away from one another. Therefore the balloon would eventually expand so large it could pop. This is evident when one puts a soda bottle in the freezer for an extended period of time. The bottle will burst due to the pushing out of the molecules."

(Unfortunately, the student used the example of water as an example for what happens when a substance gets cold.)

"The balloon got larger because it expanded. Now if the balloon was deflated and tied at the end, it would stay the same. If there water in the balloon, I assume it would get smaller."

8% **Blank**

*** *very good answer*
** *good answer*
* *minimally correct answer*

Comments

Almost half the students (42%) were able to respond with some degree of success to this question. Almost a quarter (24%) answered that the balloon got smaller but did not explain their answer or gave an incorrect explanation.

This type of response points to the advantage that open-ended testing has over multiple choice testing where the selection of the option "smaller" may have been considered correct. Here we do not give credit to those responses that did not include an explanation or gave an incorrect explanation, and we see that it is a sizable number of students who respond in that manner. The purpose of this question was to go beyond the one-word answer to get at students' conceptual understanding. The responses to this question are similar to a Grade Four question in which students were asked to explain how water could create a canyon, and 22% responded with the word *erosion* with no explanation or an incorrect one.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 8

Here is how a scientist often works:

The scientist observes something happening.

The scientist then thinks of a hypothesis or possible explanation of why or how it happened.

The scientist designs or plans an experiment to test the hypothesis to see if it is likely to be true.

When conducting the experiment, the scientist gathers data or information which the scientist can then analyze to help draw a conclusion about the hypothesis.

Suppose you made the observation stated below. In the spaces below the observation, give a reasonable hypothesis that might explain the observation. Then describe an experiment you would conduct to test you hypothesis. Tell what data you would gather if you were to actually conduct your experiment.

Observation: At the end of the marking period, the students in Mr. Smith's eighth grade English class got lower grades than students in Mrs. Jones' eighth grade English class.

Major Concepts/Abilities Tested

Process skills

Scientific inquiry

Experimental design

Use of controls

Centrality of evidence

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas, Elizabeth Badger, and Patrick Markham.

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

***** 14% Hypothesis is clearly stated and feasible; experiment addresses and will test hypothesis; relevant variables controlled (if necessary); explicit and clear description of experiment.**

Hypothesis

"Mr. Smith gives a much more difficult test than Mrs. Jones, but her tests may be a little too easy. And when the grades at the end of the term come out, the two classes have very different grades."

Design of Experiment

"Give the teachers the same information to teach and then have them make up tests on that information. Each teacher should give their tests to the other class so that Mrs. Jones' class is taking the test that Mr. Smith wrote and Mr. Smith's class is taking Mrs. Jones' test. Also have the two classes take a third test that someone else made up on the material so that we can see how they do on another test. If Mr. Smith's class gets higher grades with Mrs. Jones test, and if Mrs. Jones class gets lower grades than Mr. Smith's then the hypothesis is right. If the grades are still like that then they should move onto a different idea."

Data To Be Gathered

"Grades for Mr. Smith's class and the grades for Mrs. Jones class before they switched the tests.

The grades for Mr. Smith's and Mrs. Jones' class on the tests they switched.

The grades on the other test they took."

(Unlike most of the responses, this student was able to answer with a quantifiable hypothesis.)

Hypothesis

"Mr. Smith is not as good a teacher as Mrs. Jones.

Design of Experiment

"Give both teachers a certain amount of time, say two weeks, to teach their class about a certain time period in English, one that neither class has ever studied. At the end of two weeks, give both classes the same test on the time period. You might also give each teacher a questionnaire to see how he/she feels about his/her career."

Data To Be Gathered

"A. How many people are in each class.

B. How long each teacher has been teaching.

C. The test grades of each class.

D. The averages of everyone in the class for the past few years.

E. Review the results of the questionnaire, maybe Mr. Smith does not enjoy being a teacher."

(Although the student is on shaky ground with a hypothesis that is dependent upon a value judgement i.e., what constitutes a good teacher; this student controls for the effect of prior knowledge and the level of difficulty of the test. The questionnaire gives attitudinal information that cannot be controlled for but which might explain the results.)

**** 27% Hypothesis is less clear/feasible, but experiment addresses an appropriate hypothesis; minor weaknesses in instrumentation or control of variables.**

Hypothesis

"Mrs. Jones is a better teacher."

Design of Experiment

"About 5 students are switched from Mr. Smith's class to Mrs. Jones' class and vice versa. The students spend one quarter of the year in their new class and are told to try their hardest. Then they are switched back for another quarter of a year and are again told to try their hardest."

Data To Be Gathered

"Compare the marks of the ten students in each class. If all ten have tried their hardest and passed in the same quality of work and still receive different grades, then Mr. Smith is a harder grader. If the students try their hardest and pass in lower quality of work in Mr. Smith' class, then maybe Mr. Smith is not as good a teacher or is harder."

(This response is very similar to the previous response but the student does not account for the abilities of the students. There is a great emphasis on the intangible of "try their hardest.")

* 27% **Somewhat vague experiment; seems to address hypothesis; does not necessarily account for some variables.**

Hypothesis

"Mr. Smith's English class isn't as smart as Mrs. Jones English class."

Design of the Experiment

"I would give both classes about one week's worth of the same work and see which class would get the better grades."

Data To Be Gathered

"I would gather the work that they would do. Make them do it and get an average of their grades when is was over and see who was smarter."

(Compare this response with the response in the first category where the student collected the class averages as a basis for comparison. This student does not account for differences between the teachers or that grades are a function of the teacher, nor does the student define the word smart. The only item controlled is the material covered during the fixed period of time.)

*** *very good answer*
** *good answer*
* *minimally correct response*

Incorrect Answers

21% **Hypothesis weak or not testable; experiment does not address hypothesis or addresses it poorly; collects poor information; or no real experiment or investigation proposed.**

Hypothesis

"The students are not interested in the class."

Design of the Experiment

"Ask questions to teachers and students about homework and classwork, etc. and ask if students were participating."

Data To Be Gathered

"Answers to above questions and more hypotheses."

(This response does not generate a testable hypothesis nor does it control for any variables that might cause differences in the grades of the class nor does it measure interest except by questioning.)

3% Gives recommendations for improving test scores.

Hypothesis

"They weren't trying as hard."

Design of Experiment

"Let them come after school to get extra help."

Data To Be Gathered

"See how much they have increased their grades."

3% Other incorrect response.

4% Blank

Comments

Sixty-eight percent of the students were able to answer this question with some success indicating that, by grade eight, students have a working knowledge of how to generate a hypothesis, how to devise an experiment to measure it, and how to recognize and prioritize the data collected. This was a difficult question because of the intangibles in dealing with human behavior. Many students seemed to recognize this and dealt with it appropriately. This may be considered more of a social science type of experiment because the difference in perspective does not lend itself to traditional scientific analysis. Wherever it is categorized, the same requirements for experimental design would apply. Because a social scientist observes different phenomena, the rules of measurement don't change; only the instruments change.

One of the most important aspects of science that students learn in their early and middle years is accurate measurement. When deciding what data is important to gather, students must also decide how to measure it accurately, how to check those measurements and how to evaluate the results of those measurements. This collection and use of evidence is the basis of all scientific investigation. At the early grades, students can measure the concrete. In the middle years, students can begin to deal with more ambiguous situations in preparation for the needs of higher level science.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 8

The following speakers' arguments are about total gun control. This policy would make the possession and use of guns of any type illegal except by law enforcement of military personnel doing their jobs. Read the questions below; then read the arguments to answer the questions.

Statements On Gun Control

Speaker 1: *Many deaths result from legally-owned guns easily getting into the hands of careless children, jealous lovers, or irate relatives. Furthermore, most illegally-possessed guns were once legally owned by someone else. The control of gun sales is not enough. While it is true that making all guns illegal would mean that only criminals would have guns, there would be a lot fewer guns around for criminals to get hold of.*

Speaker 2: *Universal gun control, as I understand it, would make hunting illegal. Hunting has been a popular and respected sport as well as an essential means of securing food throughout the history of man. Also, hunting is an effective way of controlling the population of some species. Otherwise, they might suffer from mass starvation due to overpopulation.*

Speaker 3: *While the flesh of game, hunted and killed with guns, is sometimes used for food, this is not out of necessity. In fact, the money paid for a hunting license would probably buy more meat from the supermarket than is taken from the hunter's catch.*

Speaker 4: *While hunting can help control overpopulated species, this cannot be a primary purpose for it. There are many more humane methods of putting animals to death. The majority of animals shot by hunters are only wounded and suffer excruciating pain for hours before Mother Nature mercifully relieves them of their suffering.*

Speaker 5: *Citizens of the United States have a constitutional right to keep and bear arms. Our forefathers recognized the importance of guns for personal protection, food and sport. Universal gun control is unconstitutional. Besides, even if guns were made illegal, there would still be problems. Even though many drugs are illegal, their trafficking and abuse are bigger problems than ever.*

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

***Speaker 6:** An amendment to the Constitution has been mistakenly used as an argument against gun control. The Constitution does not simply state that citizens have the right to keep and bear arms. Rather, it states that citizens have the right to keep and bear arms for the purposes of maintaining a militia. The Constitution was written at a time when there was no army or national guard that could distribute weapons before military action. Fighters in the militia were volunteers who had to use their own weapons. Actually, this is a case in which the Constitution does not have to be reinterpreted because of changing times. The amendment is quite clear in its wording and intent. With the militia system no longer in use, it cannot be used to argue against gun control.*

Major Abilities/Concepts Tested

Evaluation of argument

Identifying types of information

Distinguishing fact from opinion

The percentage of student responses and appropriate examples are listed under each of the categories below.

Which speakers are *in favor of* total gun control?

4 out of 4 correct - 23%
3 out of 4 correct - 18%
2 out of 4 correct - 12%
1 out of 4 correct - 21%
0 out of 4 correct - 21%
Blank - 3%

This first question simply measures how carefully students read the opinions and whether they were able to understand what the speakers were saying. Some may argue that the phrase *total gun control* may confound students, but the introductory paragraph does state the speakers are giving their opinions on total gun control. Here we simply wanted students to read the arguments carefully in order to prepare them for the next questions. Students seemed to be fairly adept at pulling out the arguments for gun control.

Which speaker gave the strongest argument? Explain why the argument is strong.

Correct Answers

**** 11% Good explanation of strength of argument with correct interpretation of speaker's argument.**

"I think that Speaker 6 gave the strongest argument. First of all, the speaker has proof of why they think that gun control should be controlled and illegalized. Second of all, the speaker also has a proof to jump back to so that no other speaker could use the Constitution as a reason to argue back against them. From my point of view, being able to back yourself up and also backing yourself up against another argument is the strongest argument."

(This student correctly addresses the rhetorical points of the argument.)

*** 29% Provides some acceptable reason for choice of argument.**

"Speaker 6 gave the strongest argument. The reason that speaker 6 gave the strongest argument is because he or she used the American Constitution to back himself up. No other speaker used any evidence to back themselves up. They were all opinions. Speaker 6 told what the constitution told about gun control. And the Constitution is how the American people should live by. The U. S. constitution has laws written down about how the American people should live when it comes to gun control. People misunderstand the Amendment about this so it doesn't work to its fullest but I still feel that speaker 6 had the strongest argument because of his data and facts he brought along in his speech."

** *good answer*
* *minimally correct answer*

Incorrect Answers

31% Makes selection but does not provide enough explanation to determine strength (repeats speaker's argument).

"Speaker 1 because he brings up the point of many children being accidentally shot and killed as a result of many available firearms. Also that it would cut down on a lot of unnecessary crime in larger cities and innocent people being shot and killed."

6% Misinterprets speaker's statement.

"Speaker 6 gave the strongest argument because he thinks it's right to have a gun and its an amendment freedom to bear arms."

8% Ignores strength of speaker's argument; explanation relies on student's personal viewpoint.

"Speaker 1 has the strongest argument. I think he does because what he said is very true and it is reality. Many people have been killed by accidental shootings. There can be more done about shootings. We can control gun sales, we can stop crimes, and shooting. Many people are afraid to walk in streets any more because of the poor gun control. Take Boston for example, there has been over thirty murders in the first two months of 1990. What a way to bring in the new year with a bang. I think that the constitution said that guns are allowed only in the military army. If I were Mike Dukakis, I would stop the sale of guns and make them illegal. The future of the U.S. depends on it."

7% Other incorrect response.

7% Blank

Forty-three percent stated that speaker 6 had the strongest argument. Twenty-nine percent felt speaker 1 had the strongest argument, and sixteen percent felt speaker 5 presented the strongest case.

Comments

Here we wanted to know what students believed was the strongest argument and why. The important part of this question is the "why." Most students were able to select an argument they felt was strongest, but had difficulty in explaining why it was strong. Many students were unable to separate their own viewpoints from the speakers'. For example, "Speaker number 1 was the strongest because it's about killing and killing is wrong." Students who answered correctly were able to take an external stance and evaluate the argument objectively, i.e., look at the type of argument and its construction. For example, "the argument was strong because the statement was presented clearly and it anticipated arguments and answered them." Many students simply reiterated the argument as proof of its strength.

Which speaker gave the *weakest* argument? Explain why the argument is weak.

Correct Answers

** 11% **Good explanation of weakness of argument with correct interpretation of speaker's argument.**

"Speaker 3 gave the weakest argument. He/she told nothing about why we should or shouldn't have guns. Besides being hard to understand, the answer was obviously not well thought out."

(This student addressed the construction of the argument rather than the argument itself.)

* 33% **Provides acceptable reason for choice of statement.**

"Speaker 3 gave the weakest. He did not take a stand on either issue and the fact that their payment for hunting licenses would buy more meat than they would get hunting. He talked about things that weren't important and didn't make a good argument."

Incorrect Answers

19% **Makes selection but does not provide enough explanation to determine weakness.**

"Speaker 3 because he said with the money he can go to the store and buy meat instead of getting a license."

8% **Misinterpretation of speaker's statement.**

"The speaker that gave the weakest argument was speaker 6. He or she had the weakest argument because all that person was talking about was protecting themselves with weapons and defense if we get into a war."

11% **Ignores weakness of speaker's argument; explanation relies predominantly on student's personal viewpoint.**

"Speaker 2. Overpopulation of animals is a poor excuse for not making guns illegal. It is not as if deer, moose and bears are running wild everywhere. This was not a problems before guns were invented and it is not a problem now.."

8% **Other incorrect response.**

8% **Blank**

Forty-two percent felt that speaker 3 had the weakest argument, twenty-four percent felt that speaker 2 was weakest, and thirteen percent felt that speaker 4 gave the weakest argument.

Comments

This question was a little easier for students who are much more adept at describing what is lacking in an argument rather than what it provides. Students were skilled at finding fault with the logic of an argument. They seemed to believe that if they disagreed with the argument, the argument was weak, e.g., "Speaker 2 is so fake. I believe in humane methods of controlling the population, not by guns. And animals deserve to live as much as we do. We are sometimes crazy, look at this, we almost destroyed the population of gorillas and now we're killing the elephants."

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 8

"Although I worked hard in the fields, I still could not earn enough money to feed and clothe my children. I borrowed money from a merchant in exchange for a promise to work for him in his shops in the new world for seven years. My family and I packed a few belongings in trunks and bundles, said good-bye to our friends, and went to the local seaport where I signed agreement papers. There were many other men about my age at the boat docks. Many had brought their families with them. We were crowded together in the space below the ship's deck where the freight was usually hauled. The voyage across the Atlantic took over six weeks and several people died before we reached land. Finally, we set foot on soil and our new life began."

What was the speaker, and approximately when did he live? What clues in the passage lead you to your answers?

Major Concepts/Abilities Tested

Analysis and evaluation of material including weighing evidence and making inferences

Sense of history

Specifics of history

The scoring of this question was divided into four parts:

- identification of the speaker
- identification of the time period
- clues that identify the speaker
- clues that identify the time period

The percentage of student responses and appropriate examples are listed under each of the categories below.

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

Identification of the Speaker

Correct Answers

**** 51% Correct primary identification — indentured servant.**

*** 12% Correct secondary identification — immigrant, farmer, father.**

*** good answer*
** minimally correct answer*

Incorrect Answers

17% Slave.

10% Other incorrect response such as merchant, sailor, or pilgrim.

10% Blank

Identification of the Time Period

Correct Answers

**** 53% Correct response — within the early 1600s to the end of the 1700s; during the settlement of America/New World.**

*** 9% Correct response — early to mid-1800s range.**

*** good answer*
** minimally correct answer*

Incorrect Answers

12% Time too early (1600 or before).

6% Time too late (late 1800s - 1900s range).

5% Other incorrect response

11% Blank

Clues to Identity

Among the clues are the speaker's choice to leave his homeland and/or to trade passage for an agreed period of service.

Correct Answers

**** 46% Appropriate clues for correct identification.**

"The speaker above was an indentured servant. I got the idea of him being an indentured servant because he borrowed money from someone and promised to work for him in the New World for a certain period of time."

*** 9% Logical evidence for secondary identification.**

"The speaker above was a hardworking man that couldn't pay off his debts. He probably lived in the late 1600s early 1700s over in England. He promised to work for a man for seven years in the New World for his trip over to the New World and freedom to support his own family in seven years. This was a very typical thing to do during the times of beginning in the colonies. This man figured he would have a better chance of living nicely there than his present home."

(This student skirted around the proper identification of indentured servant and never really properly defined it. The definition of the speaker as a "hardworking man" was the only actual answer the student gave. Although the clues were appropriate for an indentured servant, the student never gave an indication that the answer "indentured servant" was what she was getting at. The evidence was also appropriate for the identification of a "hardworking man.")

** *good answer*
* *minimally correct answer*

Incorrect Answers

13% Evidence for incorrect identification.

"The speaker was a slave and he lived approximately in the 1800s somewhere. The clues were that he worked hard in the fields and did not make enough money so he went for a voyage across the Atlantic that took six weeks to start a new life."

(This response was typical of those who identified the speaker as a slave because he worked in the fields. Although the speaker states that he didn't make enough money to feed his family, many took that to mean it was an example of slave labor.)

13% Other incorrect response

18% Blank

Clues to Time

Among the clues to the time are: length/type of travel; others immigrating in the same manner; use of term "New World."

Correct Answers

**** 46% One or more correct clues to time.**

"The speaker probably lived around 1700. Some clues are that it took so long to travel and the man worked seven years for food and shelter and people don't do that any more and many people died on the trip over."

"The clues in this passage that lead me to my answers were that he was working a field, not many people do that now. People don't borrow money from merchants anymore, mostly banks. That was done in the late 1600s to the early 1700s. Another reason for me believing that it was early in the 1700s is because the boat trip took six weeks in order to cross the Atlantic, now it only takes a week or two at the most. Probably they would fly rather than take a ship. Also many of these things were typical of the early settlers in the American Colonies. Most of the people that came here had a hard life where they came from."

** *good answer*

Incorrect Answers

13% Logical evidence for incorrect time placement.

"This speaker probably lived in the 1500s. I got the idea of him living in the 1500s because that's when the New World started."

(Although this student did use logical evidence for his time placement, he did not use enough of the clues to deduce that although the New World may have started in the 1500s, it wasn't settled until the 1600s.)

11% Other incorrect response

"I think the speaker above was a Pilgrim who came from England to begin new colonies in the New World (America). This man must have lived in 1776 when the Mayflower left on its voyage. The clues that made me believe he was a pilgrim are: he would work in the New World. I knew this meant America. The passengers were crowded below the deck. The voyage across the Atlantic took six weeks. I know the pilgrims sailed across the Atlantic and it took a long time. Some died before we reached land. I knew the Mayflower was a hard journey."

(This was a problematic response because the student gave a date that was in the appropriate range, and used the clues —e.g., the New World and travel conditions — necessary to arrive at that date. However, the conclusions drawn from the evidence were so inappropriate that the student received credit for the correct time, but did not receive credit for her selection of clues. This response does point to why open-ended questions are so valuable. Had the student simply listed 1776 as the time, she would have been given full credit for understanding the evidence.)

29% Blank

(In the identification portion of this question, about ten percent of the students did not answer. Yet in the selection of evidence, almost twenty percent of the students could not select any clues that identified the speaker nor could almost thirty percent offer clues that identified the time period. This type of question caused difficulty for some students who felt at ease identifying the speaker and the time period.)

Comments

The key to using evidence is recognizing what evidence is relevant. Generally, students who correctly identified the speaker as an indentured servant did so because they used the right evidence. When students answered incorrectly, they generally selected evidence that did not discriminate among many different identities, but could be applied to several, such as "worked hard in the fields" or "could not earn enough money to feed my family." These clues aren't descriptive enough to exclude many alternatives.

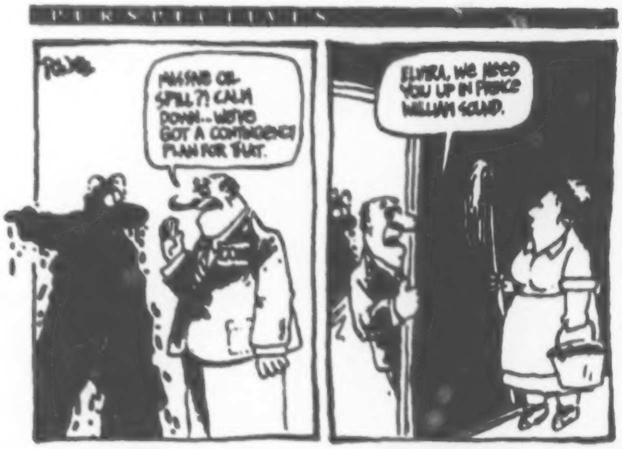
Students had to use more than one clue to come up with the correct time period. For example, students who cited only the term "New World" often gave a time too early because they only considered the exploration of the New World (i.e., the 1500s). They did not consider mention of shops in the New World as an indication of settlement (i.e., after 1620). Those who selected a date that was considered too late often referred to the crowd of people in the boat as proof of the wave of immigrants that came in the last half of the nineteenth century. But citing both the term "New World" and the crowded passage across the Atlantic would have indicated a time in which settlement had begun in the New World, before it was established as the United States of America.

These results point to two generalizations. First, students have a hard time selecting appropriate evidence. Second, students have a poor grasp of chronology of events.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 8


PERSPECTIVES
MASSIVE OIL SPILL?! CALM DOWN... WE'VE GOT A CONTINGENCY PLAN FOR THAT.
ELVIRA, WE NEED YOU UP IN PRINCE WILLIAM SOUND.

Cartoons often make fun of people and, at the same time, pass on a bigger message or idea. Who is the cartoon making fun of, and what is the main idea of the cartoon?

Identify who the cartoon is making fun of.

The percentage of student responses and appropriate examples are listed under each of the categories below.

Major Concepts/Abilities Tested

Contemporary history

Analysis of graphic representation

Correct Answers

* 44% **Primary correct/reasonable identification.**

"The cartoon is making fun of Exxon Oil Company."

* *correct response*

Incorrect Answers

21% **Incorrect identification related to event.**

"The cartoon is making fun of the Exxon oil spill in Prince William Sound."

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

(This answer is close, but not correct. The cartoon refers to Exxon's response to the oil spill, not the oil spill itself.)

25% **Incorrect identification related to figures in cartoon.**

"The dirty guy. The man is talking sophisticatedly about the maid's job."

5% **Other incorrect response.**

5% **Blank**

Identify the main idea of the cartoon.

Correct Answers

** **34%** **Primary correct identification of main idea of cartoon.**

"The cartoon is making fun of Exxon Company. It's after an oil spill and they are supposed to be prepared for any type of an emergency but the only back-up they have is a couple of people not really qualified to handle this emergency (represented by Elvira)."

* **21%** **Vague response, or one that touches upon the issue but is too specific (e.g., the cartoon is about the Alaskan oil spill, protecting the environment).**

"The cartoon is making fun of the Exxon Corporation because of the oil spill. The main idea is that there was so much damage that any attempts to aid the problem seem futile. Also, I think that the artist feels that people should pay more attention to the environment for generations to come."

(The response does not address the issue of Exxon's preparedness to face such a situation, and there is little indication of the artist's greater intent.)

** *good response*
* *minimally acceptable response*

Incorrect Answers

8% **Lack of prior knowledge/understanding of specific situation.**

"The cartoon is making fun of the president of the Exxon oil company. The main idea is how they clean up the oil spill. They are showing what a big problem cleaning up the oil can be."

13% **Literal response or misinterpretation of the cartoon.**

"There was a 'massive oil spill' and they needed someone to clean it up so they call Elvira."

15% **Other incorrect response.**

6% **Blank**

Comments

Realizing that satire and political humor are very difficult for young people to understand, the question is divided into two parts to make it easier. By getting students first to consider *who* is involved, they are alerted to the particulars that should explain the larger issues. If they were able to identify who the cartoon was making fun of, they were generally able to give the main idea of the cartoon.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 8

A community has found that large quantities of hazardous chemical wastes are being left in its dump. The EPA has determined that these wastes are being produced by a large manufacturing plant within town limits. The EPA, a state agency and local officials are working together to find a solution to the problem. What are three possible courses of action? For each course of action, what are some negative impacts?

Major Concepts/Abilities Tested

Human geography

Environmental issues

Economics

Cause and effect

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

Course of Action				
	#1	*#2*	*#3*	
**	26%	20%	18%	**Reasonable, well-developed answer where the negative impact is appropriate and well matched to the course of action; more than one acceptable impact.**

Course of Action

"Shut down the manufacturing plant."

Negative Impacts

"Some negative impacts are that there would be higher taxes because the town isn't manufacturing as much and not getting as much money in taxes from the industry. Another negative impact is that all the people who worked at the plant will be out of jobs."

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

*** 33% 32% 27% Reasonable but less well-developed answer.**

Course of Action

"Place a fine on dumping."

Negative Impacts

"The company would raise its prices on its products to pay for the fine."

** *good answer*
* *minimally correct response*

Incorrect Answers

11% 9% 7% Vague answer, minimal development.

Course of Action

"Dispose of it in other towns."

Negative Impacts

"People might not want or wouldn't want chemicals in their community."

(There is a fine line between the second and third answers, but the second answer shows some thought as to the direct impact of the action. That "people might not like it" is a negative attitude, not a negative impact.)

5% 5% 4% Reasonable action with unreasonable illogical impact.

Course of Action

"They could close the dump."

Negative Impacts

"Then no one will dump hazardous waste there."

(The impact is not considered negative.)

17% 20% 22% Other incorrect answer.

Course of Action

"Stop waste."

Negative Impacts

"All people should stop all waste."

(The students restated the course of action as a negative impact.)

9% 12% 20% Blank

Comments

Generally students were very reasonable in their suggested actions. It's clear, though, that they had trouble coming up with three separate courses of action. Students recognized the complexity of the problem, as well as the need to consider more than one side of the issue. Although they appeared to be very attuned to the environment, they were able to evaluate economic issues with some strength. Seldom did they produce extravagant responses. And when they did, often the negative impact showed that the student understood its extravagance. For example, many students who gave as a course of action "shooting the waste into space," discussed such negative impacts as "the high cost, and the damage the waste could inflict on space." What may have appeared frivolous initially was actually well thought-out.

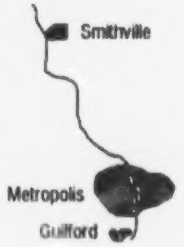
Students also realized that a quick fix was not a real solution, but merely delays the need for a solution. They realized that there is often a trade-off between environment and economics, and people often live with some negative environmental factors in order to achieve economic stability. Most realized that, despite the plant being the source of the pollution, it would be better to work with the plant to find a solution rather than to close it.

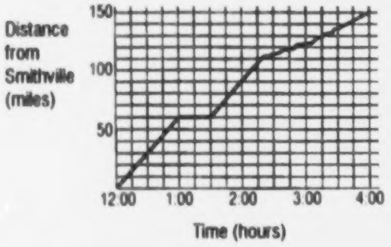
ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 8

The Mitchells took a 4-hour car trip from Smithville to Guilford. The map and graph below help describe their trip.





Use the map and the graph to describe what the Mitchells were doing during each 1-hour interval. Tell as much as you can about the trip (e.g., kinds of roads, traffic jams, stops, etc.)

Major Concepts/Abilities Tested

Map reading

Graph reading

Ability to relate information from different sources

The percentage of student responses and appropriate examples are listed under each of the categories below.

12:00 - 1:00

Correct Answers/Methods

*** 10% **More than one legitimate inference, but must include rate of travel.**

"From 12:00 to 1:00, the Mitchells drove 60 miles. They made good time going about 60 miles an hour. No traffic jams, roads look straight for the most part."

"The Mitchells travelled at a constant speed of 60 miles per hour. Because they were traveling at this speed, they were most likely driving on a highway."

** 46% **At least one legitimate inference.**

"They drove 60 miles. Didn't have to make any stops and the roads were clear.

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

They didn't seem to run into any traffic jams because of accidents or detours."

* 14% **Minimal response or some inaccuracies.**

"Between 12:00 and 1:00, the Mitchells traveled 60 miles."

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers/Methods

10% **Attributes characteristics of graph or map to road.**

"They traveled on a straight road, going northeast. They traveled 51 miles."

15% **Other incorrect answer.**

"They left Smithville. Probably stopped at a restaurant for some lunch and then headed off again."

4% **Blank**

Comments

Students were most successful in interpreting this part of the Smith's journey. The positive and uninterrupted slope of the line suggests highway travel, as does the fact that the road skirts Smithville without entering the town. Although only a small proportion of the students explicitly stated the resulting rate of 60 miles an hour, the graph depicted a steady rate of speed, and the distance traveled was easy to read.

1:00 - 2:00

Correct Answers/Methods

*** 9% **More than one legitimate inference, but must account for complete stop and starting again at same rate.**

"From 1:00 to 1:30 they stopped driving, probably to get a bite to eat and some gas for the car. At 1:30 they again started to travel at a constant speed of 60 miles per hour."

** 29% **At least one legitimate inference.**

"They were still driving. They might have stopped for gas and at a restaurant to get a drink and use the rest rooms."

* 24% **Minimal response or some inaccuracies.**

"During this hour they traveled 30 miles. They probably ran into a lot of traffic jams."

Incorrect Answers/Methods

11% **Attributes characteristics of graph or map to road.**

"I think that they hit a detour or a place where they weren't supposed to go such as a forest because they changed course."

"They have traveled about 60 miles and have seemed to have gone steady to the right."

24% **Other incorrect answer.**

"They had traffic between 1:00 and 1:30. They didn't get much distance, but for the last half hour they cruised down the highway and drove another 60 miles in a half hour."

4% Blank

Comments

The same proportion of students who had correctly inferred that the Mitchells started their journey by traveling on a highway at a rate of 60 mph, recognized that they came to a complete halt from 1:00 to 1:30, then continued their journey at the same rate as before. However, most students found the change in the direction of the graph difficult to interpret. Although some students suggested a stop, few appeared to recognize that the graph depicted the relationship between the two variables of time and distance and that the horizontal line signified that one (distance) did not increase in relation to an increase in the other (time).

2:00 - 3:00

Correct Answers/Methods

***** 9% More than one legitimate inference, but must include arrival in city.**

"They went through Metropolis. They probably got in a few traffic jams because it is a large city with, probably, a lot of highways. These traffic jams would not be bad though, and would clear out rather quickly."

**** 25% At least one legitimate inference.**

"The Mitchells started out traveling at a constant speed but seemed to get themselves into some sort of a traffic jam."

"From 2:00 to 2:15 they were all right, but from 2:20 to 3:00 they were on an ugly road with loads of traffic."

*** 21% Minimal response or some inaccuracies.**

"About 30 miles, hilly roads, stops."

"At 2:00 they had gone 90 miles. At 2:30 they had gone 111 miles. At 3:00 they had gone 122 miles."

Incorrect Answers/Methods

27% Attributes characteristics of graph or map to road.

"The Mitchells were driving at a fast moving pace despite the steepness of the ascent and then the rugged steepness."

"From 2:00 to 3:00, it started out pretty straight. Then they must have gotten onto a secondary or a dirt road because things were pretty bumpy for them."

14% Other incorrect answer.

"During the third hour, they probably stayed on the same road with no stops. Also no traffic."

"They are 14 miles away now. They are traveling through a meadow near the seas or lakes."

5% Blank

Comments

In this section of the trip, the map contributes a great deal to understanding what was happening during the journey since the Mitchells' arrival at the city was, undoubtedly, the cause for their unsteady rate of travel. Most students appeared to ignore the map, concentrating on the information contained in the graph. About a quarter of the students recognized that the variability of the slope reflected stop and go traffic, but many concentrated only on reading the graph (e.g., number of miles traveled) rather than interpreting it. The variability of the graph also led many students to interpret the slope as a depiction of the road itself. In this section, twice as many students as before attributed characteristics of the graph to the road, forgetting that the graph represented the rate of travel, not its conditions.

3:00 - 4:00

Correct Answers/Methods

*** 6% **More than one legitimate inference, but must include leaving the city.**

"For this 1/2 hour, they were still in Metropolis. They traveled 15 miles. The second 1/2 hour, they were out of Metropolis and traveled the last 15 miles to Guilford."

** 26% **At least one legitimate inference.**

"For the first half hour it was stop and go, but then it picked up and the last half hour was smooth driving."

* 24% **Minimal response.**

"It took 150 miles but the Mitchells have finally reached Guilford after 4 hours of travel."

Incorrect Answers/Methods

21% **Attributes characteristics of graph or map to road.**

"At about three-thirty, the bumpy road turned back into the smooth ride they had earlier."

18% **Other incorrect answer.**

"More traffic jams and there were holes in the road. Some main highways were blocked."

5% **Blank**

Comments

Students' responses to this section were similar in kind to the previous section. Those who were able to use both sources of information noted that the Mitchells spent much of the hour navigating traffic through Metropolis, but for the last 20 minutes of their journey they were again on an open-road, though travelling at a much slower rate than they had done on the highway.

In general, this proved to be a difficult problem for students. Students who understood how to read the graph performed in a similar manner over all four intervals. Those who did not understand varied in the extent to which they attributed characteristics of the graph or map to the road. Many, undoubtedly, did not understand what the slope of the graph signified. Others found difficulty in integrating two different types of information.

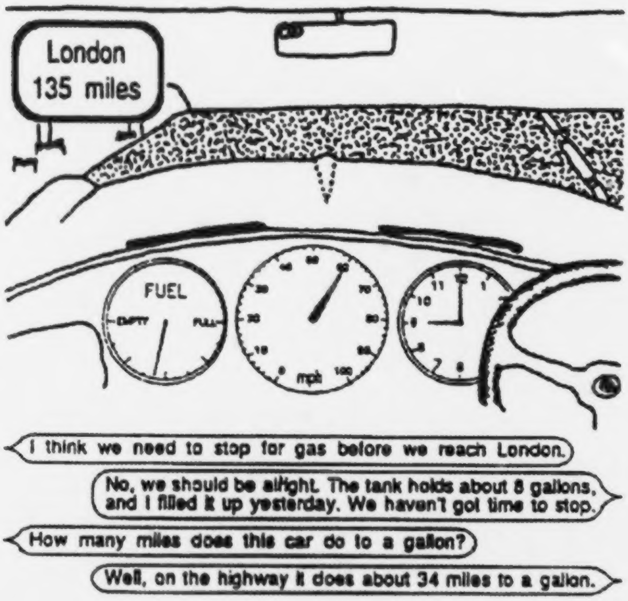
This question was adapted from an instructional unit on informal use of functions which was developed by the Shell Centre for Mathematical Education. Its assumption is that a familiarity and confidence in using functional graphs allows students to develop a deeper understanding of the kind of relationships that underlie mathematics.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 8

The conversation below is between a driver and a passenger on their way to London by car. The picture shows what they see as they speak.


London
135 miles
FUEL
EMPTY
FULL
mph
I think we need to stop for gas before we reach London.
No, we should be alright. The tank holds about 8 gallons, and I filled it up yesterday. We haven't got time to stop.
How many miles does this car do to a gallon?
Well, on the highway it does about 34 miles to a gallon.

Do they have to stop for gas? (Explain your reasoning carefully. It will count.)

Major Concepts/Abilities Tested

Recognizing the mathematical structure in an unstructured problem

Interpreting graphical data

Identifying relevant information

Integrating information

Computational facility

The percentage of student responses and appropriate examples are listed under each of the categories below.

* Adapted from a problem developed by the Shell Centre for Mathematical Education.

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

Correct Answers/Methods

** 22% **Yes, by computing how far the car can go on the amount of gas in the tank (approximately 119 miles).**

** 4% **Yes, by computing the number of gallons needed to drive 135 miles (3.9 gallons)**

* 9% **Yes, by approximating appropriately.**

"I think they should stop. They might have enough, but it would come close because they have less than 4 gallons left."

** *good answer*
* *acceptable answer*

Incorrect Answers/Methods

21% **Yes; incorrect or no reason given.**

"Yes, they do need to stop for gas because they're going 60 MPH, and they're going to run out of gas soon."

"Yes, because they only get 34 miles to the gallon."

10% **No; assumed car had full 1/2 tank (4 gallons).**

18% **No; assumed car had full tank (8 gallons).**

15% **Other incorrect response.**

"No, they probably don't have to stop because London is 135 miles and the car does about 34 miles per gallon. They would use up about 3 gallons and have extra gas left over."

1% **Blank**

Comments

This question challenged students to identify and coordinate different kinds of information in order to come to a judgment. In contrast to most math "problems" where the consequences of a wrong answer are merely formal, an incorrect decision here could bring about unpleasant real-life consequences. As such, it is typical of the kinds of situations in which we typically use mathematics and it requires the same kinds of thinking skills. There is an abundance of different kinds of information, from different sources. Some of the information is necessary in order to solve the problem; some is redundant; some is irrelevant. An adequate solution comes, not from the calculations required, but from the ability to translate the situation into some simpler structure. In this case:

Given that a car uses 34 miles to a gallon,

How far will it go on *xx* gallons?

or

How many gallons will it take to go *xx* miles?

Once the structure of the problem is recognized, students are then required to interpret the information given in order to find the right values. This proved to be an obstacle for many. Although a quarter of the students were able both to recognize the mathematical structure and to integrate the various pieces of information given in the picture, approximately the same number misread the fuel gauge or did not read it at all.

It requires students to reason mathematically in a practical context. Aside from their use in evaluating students' understanding, problems like this carry a message about mathematics itself. Unlike the more typical short answer problems that have already been formulated by the teacher or textbook and are primarily confined to the manipulation of numbers or formulae, problems such as these illustrate how mathematics is used in order to structure the quantitative information that permeates our everyday existence. It shows mathematics as a way of simplifying that information in order to make reasonable decisions. This notion of how mathematics can be used often escapes students who state that what they learn in class has little utility in their lives.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 8

The percentage of student responses and appropriate examples are listed under each of the categories below.

On John's tenth birthday, John's grandfather gave him $10. He gave John $20 on his eleventh birthday and $40 on John's twelfth birthday. Following this pattern, John's grandfather plans on giving John $70 on this thirteenth birthday, but John expects $80 from his grandfather on that day. John's sister says that both amounts could be correct.

Who is right—John's grandfather, John, or John's sister?

Major Concepts/Abilities Tested

Pattern recognition

Reasoning about relationships

Correct Answers/Methods

**** 19% John's sister: recognizes both sequences.**

"John's sister. His grandfather could have given him $80. Then it would have been like doubling it each time. Or his grandfather could have given him $70 by adding $10, $20, $30 in that pattern to each year before's amount."

*** 35% John: recognizes sequence.**

"John is right, because each year his grandfather doubles his amount from the last year."

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

* 11% **Grandfather: recognizes sequence.**

"John's grandfather because every year he keeps adding in tens."

\$10 - \$20 = 10
\$20 - \$40 = 20
\$40 - \$70 = 30

** *complete response*
* *minimal response*

Incorrect Answers/Methods

25% **No recognition of any sequence.**

"I think the grandfather is the right one because he is the one giving John the money, so it's his decision."

"John is right. There is a certain pattern that his grandfather is using to give him money. 70 wouldn't fit into it."

10% **Blank**

Comments

There are three different possible sequences in the situation presented. One (recognized by John) is a doubling of the gift each year, which would result in a \$80 gift. The other two (one of which is followed by John's grandfather) would result in \$70. Whether John's grandfather arrived at this figure by adding \$10 to the increase from the previous year or by summing together the gifts of all previous years cannot be discovered until John reaches his fourteenth birthday.

While 19 percent of the students recognized at least two of the possible patterns, 35 percent seized upon the most obvious doubling pattern. Another 28 percent did not recognize any pattern and/or described an incorrect one. Five percent of the students, not recognizing the problem as a mathematical one, stated that the grandfather could do what he liked with the money.

More typical sequence tasks, which call for a single solution, can suggest that mathematical reasoning consists of finding the one correct answer. However, students should learn that a few terms in a pattern can mean very little. Discovering a pattern is basically an inductive exercise in which success is dependent upon the quantity of the data. Problems such as this promote such thinking.

Not only is the recognition of sequences and patterns important elements of inductive reasoning, it lays the groundwork for later work with mathematical functions. Exploring different numerical patterns with graphs or tables allows students an opportunity for intuitive understanding of the relationship between variables. The need to express that relationship leads naturally to expressions or equations. When students see algebraic equations as symbolic ways of describing an existing relationship between numbers, they are better able to recognize their generality.

Nature provides a rich source for the exploration of patterns, as do number sequences themselves. However, students should not only should be exposed to the patterns devised by others, but they should be encouraged to make up their own sequences. Playing such games as "What's My Rule?" builds up a confidence with using numbers, as well as the possibility of discovering interesting relationships.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 8

A very short story is going to be broken down into four parts. Read the first part; then answer the question that follows it. Only after you have answered the question that follows that part are you to go on to the next part.

Part 1

The stillness of the morning air was broken. We headed down the bay.

What are some questions you would want to have answered so that you would understand what this story is about?

Major Concepts/Abilities Tested

The ability to recognize the kinds of information needed to construct meaning from the text.

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** 68% **At least two questions that reflect experience with narratives, e.g., questions about plot, characters, setting, title, genre.**

"What are they planning on doing? Who is in the story?"

* 15% **At least one question described above.**

* 3% **Other legitimate questions.**

"What does 'the stillness of the morning air was broken' mean? What does 'headed down the bay' mean?"

"How was the morning air broken?"

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

- *** 9% Some combination of above.**

** *very good response*
* *adequate response*

Incorrect Answers

3% Irrelevant questions or questions about insignificant details.

"Was the bay pretty?"

1% Blank

Comments

Reading experts generally agree that reading involves more than decoding the author's meaning. Rather, readers construct their ideas as they progress through the text. The purpose of this set of questions was to chart that progress. The passage is taken from one employed by Stephen Norris and Linda Phillips.[1] The questions attempt to measure how well students build meaning for themselves.

This stress on construction of meaning implies activity. Readers must be aware of the kinds of information that they need and to what extent their own knowledge is relevant. Although the description of this process may appear to be abstract, in the everyday process of making sense of the written word, there are strategies that help. One is the practice of anticipating the kinds of information that will be needed to make sense of the passage. The first question, which asked students to list some questions that would help them better understand the story, attempted to measure students' ability to recognize relevant kinds of information.

Part 2

The net was hard to pull. The heavy sea and strong tide made it even more difficult for the girdle. The meshed catch encouraged us to try harder.

What do you think is going on in this story? Why do you think so?

Major Concepts/Abilities Tested

Inferential comprehension.

Correct Answers

*** 79% Correct inference based on context clues.**

"Fishermen are trying to pull a net full of fish onto the boat. I think so because they are out to sea pulling a net out of the water."

"They are fishing. The net told me they were trying to catch something and the 'heavy sea and strong time' told me that they were in water."

Incorrect Answers

16% Correct inference, no reasons given.

5% Incorrect

1% Blank

---

[1] Norris, Stephen and Linda Phillips. (1987) "Explanations of Reading Comprehension: Schema Theory and Critical Thinking Theory." *Teachers College Record.* Vol. 89, 2, 281-306.

Comments

Can students use contextual clues to construct a general meaning from a text, despite the use of unfamiliar words (girdie) and phrases (meshed catch)? The difficulty of a text can be a function of vocabulary, but often there are other factors, such as the complexity of concepts or the structure itself, that account for whether or not students understand the meaning. This question measured students ability to use the context to overcome difficult vocabulary. Almost 80 percent of the students were able to do this with justification. Another 16 percent gave a correct inference, although they were unable to state how they came to their conclusions.

Part 3

With four quintels aboard, we were now ready to leave. The skipper saw mares' tails in the north.

Explain the meaning of this part of the story.

Major Concepts/Abilities Tested

Inferring the meaning of words from context.

Correct Answers

** 30% **Offers a reasonable interpretation of quintels and mares' tails.**

"The crew had caught a certain number of fish and they were read to head in. The skipper of the boat saw whales' tails coming out of the water in the north."

"I suppose the four quintels are people or workers. Now since they were there, they were ready to go. The skipper, or boat driver, saw mares' tails to the north. Those are fish. He started heading north toward them, hoping to accomplish the task."

* 22% **Offers a reasonable interpretation of quintels only. (Of this 22%, one percent gave an accurate definition of quintels.)**

* 11% **Offers a reasonable interpretation of mares' tails only. (Of this 11%, six percent gave an accurate definition of mares' tails.)**

Incorrect Answers

14% **Focuses on "ready to leave"; ignores quintels or mares' tails.**

"The meaning of this part of the story is that now that the ship had gotten what it had come for, it was now ready to leave and get some more things. The skipper wants to get some more stuff for his luxury."

12% **Other incorrect (e.g., repeats story).**

"Four quintels were aboard a ship and now the ship was ready to leave. The skipper saw mares' tails in a northern direction."

11% **Blank.**

Comments

Although a few students knew that a quintel is a measure of weight, and mares' tails is a type of cloud formation, most students deduced the meaning of the words from the context. If their responses were reasonable within the context of the sentence (e.g., the quintels are people or cargo; the mares' tails are fish), they were credited. Those who did not attempt an interpretation of the terms were given no credit because it was impossible to explain this part of the story without making some inference about the meaning of the words. Overall, 63 percent of the students made some reasonable attempt to interpret one or both of the unfamiliar words. The others either ignored the terms or defined them in a manner that made little sense, given the setting of the story.

Part 4

We tied up to the wharf. We hastily grabbed our prongs and set to work. The splitting was done by the skipper. The boys did the cutting and gutting.

Explain in as much detail as possible what you think is happening in this part of the story. Include in your explanation what you think the words "splitting, cutting, and gutting" mean.

Major Concepts/Abilities Tested

Inferential comprehension

Correct Answers

**** 68% Cleaning or preparing the fish; explains "splitting, cutting, and gutting."**

The boat tied up to some dock. They got their equipment and started cleaning out the fish. The splitting (cutting in half) was done by the captain. The cutting (chopping off the head and tail) and gutting (cleaning out of the insides) was done by the boys.

*** 7% As above, does not explain vocabulary.**

Incorrect Answers

11% Answer analyzes context clues, but does not connect activity to preparing fish.

7% Other incorrect

6% Blank

Comment

As in the case of the previous question, approximately two-thirds of the students were able to interpret the meaning of the vocabulary and infer the activity described in the text. There was a greater tendency on the part of average and poorer readers to concentrate on the vocabulary itself, interpreting "splitting, cutting, and gutting" without any reference to the setting of the story.

General Comments

The purpose of this set of questions was to model a sequence of questions that might be asked of students as they read a narrative or, preferably, be asked by the students themselves. Obviously, the format is artificial. It is not suggested that the act of reading be continually interrupted by external questioning. However, interpretations of text cannot be used by themselves in judging reading quality because people reading equally well can arrive at different interpretations. What is important is the support that people offer for their interpretations. Consequently, it is useful for teachers to constantly check the understanding of their students, as well as for students to learn to monitor their own understanding. Misinterpretations of ambiguous phrases and vocabulary can quickly lead to larger misconceptions. If reading is a process of actively constructing meaning from the text, students should learn how to make sure that the meaning which they construct is reasonable.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 8

The citizens of Mansfield, Massachusetts are debating whether to allow a Bruce Springsteen concert in their town. Frequently, a person's argument is based on an unstated assumption—a belief the person has which affects what the person says or does.

The police chief stated that he opposed the concert because there are not enough officers on the police force.

What is an unstated assumption behind the chief's argument?

Is the assumption a reasonable one? Why or why not?

Abilities Tested

Critical analysis and evaluation

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** 42% **Answer presents a valid assumption related to the chief's argument and evaluates this assumption.**

Unstated assumption: "That the concert would be wild and there wouldn't be enough officers to restrain the crowds."

Reasonable? "Yes, because Rock n' Roll concerts tend to be wild."

Unstated assumption: "That fights break out at concerts between teenagers—since he's old and thinks all teenagers do drugs."

Reasonable? "No. It is stereotyping the people who have concerts by assuming too much of the teenagers."

Both students correctly identified the assumption behind the police chief's argument but the quality of their evaluation differed. The student who concluded that the chief's assumption

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

was well-founded referred to his knowledge of the past history of other concerts where police had been needed to restrain the crowds. The other student argued **ad hominem**, *suggesting that the police chief's assumption was based, not on facts, but on his personal perception of teenager behavior.*

* 15% **Answer presents a valid assumption but comments on the chief's argument instead of evaluating the assumption behind his argument.**

Unstated assumption: "There aren't enough officers available, if a riot breaks out many people could be injured or even killed."

Reasonable? "Yes, you shouldn't take chances when people's lives and well being are at hand."

This student accepts the assumption (that police control is necessary) as a given and proceeds to evaluate the police chief's argument, which is that there are insufficient police for the task.

** *full and complete response*
* *satisfactory response*

Incorrect Answers

23% **Answer addresses the reasonableness of the chief's argument.**

Unstated assumption: "The reason the chief opposed the concert is because of the number of officers."

Reasonable? "Yes. Safety is a big factor and if there is not a substantial amount of law enforcers, then it may be a mistake to allow the concert to take place. On the other hand, private security could be hired."

This student explains why the chief may be opposed to the concert and judges whether or not this argument is reasonable without recognizing that the validity of the argument rests on whether or not the assumption itself is reasonable.

6% **Answer accepts chief's argument and discusses how to deal with it.**

Unstated assumption: "He needs more police on this force."

Reasonable? "To get police so he can be somewhere else helping others. That is not much to ask."

10% **Other incorrect.**

Unstated assumption: "The reason he gave wasn't a real good reason, it was just an excuse."

Reasonable? "No. Bruce Springsteen is liked by many people in Mass and it just wouldn't be right to take the concert from him."

The student addresses the police chief's argument. Her answer hints at the possibility that the argument was motivated by an unstated belief, but she fails to identify that belief. Instead, she evaluates the consequence of the action and decides that it is not justified.

3% **Blank**

Comments

The ability to analyze and evaluate arguments is an important objective of the social studies curriculum; however, it is also a factor in reading comprehension. Unless students are able to identify the unstated assumptions that underlie argument, they cannot evaluate argument effectively. However, this ability to analyze and evaluate material goes beyond the critical appraisal of expository writing. It can be equally necessary in the understanding of literary characters' actions and attitudes. In all reading, it is an important aspect of comprehending.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grades 4 & 8

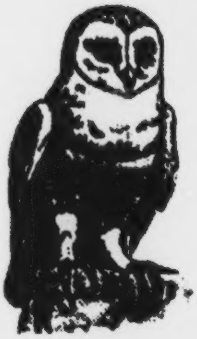
Different birds live in different areas or habitats (such as woods, lakes marshes or open areas). They eat different foods (such as fish, seeds, insects, or small animals). Birds with webbed feet swim well and live on or near water. Hummingbirds have long, thin beaks which they use to draw nectar from flowers.

Look at the pictures of the birds and notice how they are different. What do you think is the main habitat and the main food of each? Write your answers in the spaces to the right of each picture. Also explain why you chose the answer you did for each bird.

Major Concepts/Abilities Tested

Life science
Adaptation
Observation skills
Functions of characteristics
Diversity of life
Habitats

The percentage of student responses and appropriate examples are listed under each of the categories below.



Correct Answers

	Gr4	Gr8	
***	9%	14%	**Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.).**

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

"I think the reason this bird lives in a high branch is because he/she has sharp nails that grip into the tree and the bird's feet are shaped that way. I think the bird eats small rodents because it can swoop down and grab it with its sharp claws." (grade 4)

** 13% 25% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"This bird has long nails to catch the animals and it's beak is tilted inward to eat the animals better." (grade 8)

(Although the student mentions two characteristics, they both relate to food.)

* 6% 2% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"He eats rodents because look at the feet they were made to grip and lock. (grade 8)"

(Although there is a suggestion that the claws are for catching and grasping rodents, it has not been made very clear.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

51% 35% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"Owls live in the woods and come out at night for food and such. I read a book about on owls and found out about them. (grade 4)"

4% 4% **Mentions characteristics but does not relate.**

"This owl has very long, sharp, powerful claws and legs. The wings are powerful too. (grade 8)"

3% 5% **Relates physical characteristics not evident from picture to habitat and/or food.**

"This bird us dark and has very dull colors. It is its camouflage and when a victim comes along, the mouse probably scatters and the bird dives and catches him." (grade 8)

3% 5% **Uses "because it looks like" as a reason.**

"His nose looks like he would eat insects. To me he looks like a woods kind of bird." (grade 4)

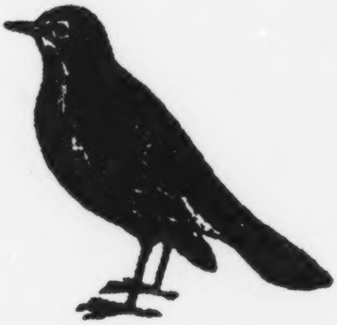
7% 4% **Other incorrect response.**

"He eats caterpillars because it shows him sitting on a tree and trees usually grow in the woods and caterpillars crawl on trees." (grade 4)

3% 8% **Blank**

Comments/Owl

It's clear from the results that students learn a lot about owls and feel compelled to share their knowledge. The owl is the only bird of the four given that elicited such a large number of students responding in this way. It could be that the owl is so interesting and unique and such a common cultural symbol that students know more discrete facts about them. It is possible that students could have incorporated prior knowledge into a correct response. Many students did. For example, if a student knows that owls eat mice, then he or she may determine that the claws are large for tearing mice simply from knowing what owls eat and observing the picture of the owl. Fewer students left this bird blank, indicating that they have some knowledge about owls. However, it seems as though those blanks were picked up in the prior knowledge category.



Correct Answers

Gr4 Gr8

*** 5% 13% **Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.)**

"The beak looks like it digs worms from the ground because of the pointed end and the feet are somewhat webbed and looks like it could walk on mud of marshes." (grade 8)

(Although the bird did not have webbed feet, the picture could been interpreted as such. The student received full credit for the response because he/she was able to make the appropriate inferences. We were not looking for correctness of response regarding what the actual bird ate and where it lived; we were assessing students' ability to observe a characteristic and make a judgement based on that characteristic.)

** 17% 24% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"They have pointed beaks to get seeds from the ground." (grade 4)

* 9% 3% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"This bird has webbed feet and there are many insects near marshes." (grade 8)

(There is no explicit relationship between webbed feet and insects; the inference being that a bird with webbed feet would live near the water, and because insects live near the water the bird would eat insects.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

28% 21% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"This bird is a robin and usually found in the wide open area; for example, your yard." (grade 8)

5% 6% **Mentions characteristics but does not relate.**

"He doesn't have claws or a short fat beak. Instead it has a long skinny beak with webbed feet." (grade 4)

8% 9% **Relates physical characteristics not evident from picture to habitat and/or food.**

"I'd say it's main habitat is in a tree because this is a rather small bird and needs a warm and safe place for itself and its babies. I'd say it's main food is worms or insects because it is a small bird and would not be able to kill or swallow mice like an owl." (grade 4)

(Although it may be prior knowledge that leads the student to believe that the bird is small, the picture does not indicate the bird's size. In the pictures, the largest bird, the heron, actually appears to be the smallest. When dealing with size, there are some size qualities that we do accept, for example, students' responses mentioning the heron's long beak are accepted as correct because, in relation to the rest of the heron's body, the beak is very long.)

8% 8% **Uses "because it looks like" as a reason.**

"This bird looks like it could be seen anywhere and live comfortably at any of the locations. It also looks like it eats worms and other insects." (grade 8)

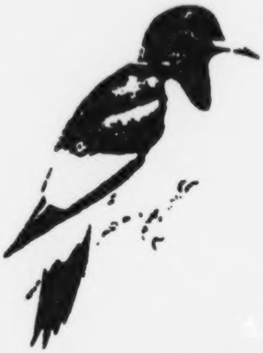
13% 5% **Other incorrect response.**

"This bird has eyes on each side of his head so he or she can see things coming toward him. That is for protection." (grade 4)

7% 11% **Blank**

Comments/Robin

About the same number of students answered this question correctly as did the previous question. Although the robin is as familiar to most students as the owl, fewer students answered in the category of prior knowledge. It could be that, without seeing the colors of the birds, students did not recognize the robin. There may have been some confusion for some students who recognized the robin, but then saw the webbed feet and could not adjust their prior knowledge of robins to what they saw in the picture.



Correct Answers

Gr4 Gr8

*** 8% 17% **Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.)**

"It probably lives in tree trunks because it can dig its claws into the side of the tree to build its nest. Also, I think it eats bugs because it has a long beak that it can put into a tree and pull out bugs." (grade 4)

** 17% 24% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"This bird's long beak looks good for reaching into the home of insects because it's long and narrow." (grade 4)

* 7% 3% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"The woodpecker likes to carve out wood from the trees with its sharp beak and finds and eats the insects." (grade 8)

(The student response is fairly clear, but the relationship between the characteristic and the food is very vague, bordering on incorrect.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

29% 20% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"The woodpecker pecks on trees in the woods to get insects out of the trees."

(This answer is very similar to the third response above, but there is no mention of the beak or its qualities.)

8% 6% **Mentions characteristics but does not relate.**

"He has sharp claws and a pointed beak." (grade 8)

7% 5% **Relates physical characteristics not evident from picture to habitat and/or food.**

"Since I have not seen a bird like this around, I suspect it hides itself in someplace where it can't be seen easily—the woods. It is a smallish bird so I would think it eats insects." (grade 8)

7% 8% **Uses "because it looks like" as a reason.**

"He looks like he is too small to eat meat." (grade 4)

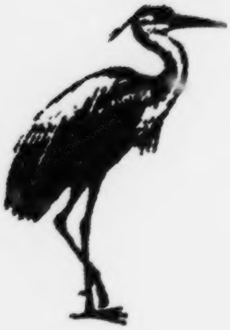
10% 5% **Other incorrect response.**

"This bird looks for a place where he can see more clearly." (grade 8)

8% 13% **Blank**

Comments/Woodpecker

Student responses here are very similar to those of the robin. It appears that the more usual looking the bird, the less students have to say about them. Compare the results of the robin and woodpecker with those of the heron which follow.



Correct Answers

Gr4 Gr8

*** 14% 34% **Relates appropriate characteristics to BOTH habitat and food. (Can be the same or different characteristics but must mention both habitat and food.)**

"The long legs can hold him above the marshes and the long pointed beak is ideal for stabbing fish." (grade 8)

** 20% 22% **Relates 1 or more physical characteristics to EITHER habitat or food.**

"This bird has long legs, a long neck and a long beak that could allow him to walk in the water and dunk his head into the water to get fish." (grade 8)

(Although three separate characteristics are mentioned, the student only discusses them in relation to food.)

* 7% 3% **Relates 1 or more physical characteristics to habitat and/or food but relationship is VAGUE.**

"They have webbed feet and if they live near or on water then they are most likely to eat fish." (grade 4)

(Here there is an inference of the relationship between webbed feet and fish as food although it is not well elaborated.)

"Because he has long legs you would probably see him near the water so he would probably eat fish." (grade 8)

(There is no indication as to why long legs would make the bird a water bird.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

29% 16% **Gives FACTS (not necessarily correct) about bird shown (prior knowledge).**

"It's a flamingo and flamingoes live in marshlands and marshlands are mostly bugs and fish." (grade 4)

8% 5% **Mentions characteristics but does not relate.**

"It has long webbed feet, a long, sharp beak and a long neck." (grade 8)

5% 1% **Relates physical characteristics not evident from picture to habitat and/or food.**

"He is big and is easily able to hunt small animals." (grade 8)

5% 2% Uses "because it looks like" as a reason.

"This bird looks like a bird from the jungle who eats fish." (grade 4)

6% 4% Other incorrect response.

"This is the ancestor of the swan." (grade 8)

8% 13% Blank

Comments/Heron

At both grade levels, students were most successful when writing about the heron. Approximately ten percent more of the fourth graders and twenty percent more of the eighth graders were able to answer this question successfully than were able to answer the other birds given; owl, robin and woodpecker. What is it then about the heron that lends itself to better responses? One answer may be that with its long legs, neck and beak, the characteristics are far more exaggerated so students are able to more easily make connections about adaptations. Another might be that the bird is far more exotic and therefore less familiar to New England students. The owl carries cultural baggage, the robin and woodpeckers are both fairly mundane, but the heron is not an everyday bird.

Comments

The responses to this question suggest that although many students are able to make inferences based on observations, a large number have problems with this type of thinking. It is a habit of mind rather than a piece of discrete information. One way of opening students to this type of thinking would be to begin with a question rather than with a fact. For example, these questions, "Why does the owl have such strong, sharp claws?" and "How do the owls' claws affect how they survive?" open the door to a more unified and systematic way of thinking and looking at something. By connecting the structure of the animal with its environment and diet, we can show the interrelatedness of all species. Students will eventually come around to the desired conclusion, but it will be through a series of observations and discoveries rather than memorization and recitation. This type of subject leads to a discussion of adaptation and how, over the generations, species carried characteristics that enabled them to survive.

Another way of getting students to examine function and form is to ask them to draw a bird that lives in the water and eats fish, as opposed to a bird that lives near the water and eats plankton. This method would direct students to forms of adaptation.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grades 8 & 12

The moon orbits the earth. Explain why the moon doesn't fall to earth.

Major Concepts/Abilities Tested

Physical science

Motion

Gravity/inertia

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

	Gr8	Gr12	
***	0%	3%	**Mention of vectors and resultant motions caused by inertia and gravity with explanation or drawing.**

"The moon has both angular and linear velocities. While the moon orbits the earth, the earth pulls down on it with the force of gravity. The moon fights the pull of gravity and wants to fly away from the earth. The moon stays in orbit and moves with a velocity tangent to the elliptical orbit (vector diagram included)." (grade 12)

	Gr8	Gr12	
**	6%	12%	**Answer discusses two dimensional motion resulting from inertia and gravity.**

"The moon is kept in orbit by the gravity of the earth and the moon wanting to go off in a straight line.

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas, Elizabeth Badger, and Patrick Markham.

The two opposing forces keep it in a circle around the earth." (grade 8)

(The student mentions two separate motions but does not indicate vectors. This response is weaker than one predicated on vector analysis but it does indicate a high degree of comprehension.)

* 10% 11% **Correctly describes either the effect of gravitational force or the inertial motion of the moon.**

"The moon is extremely distant from the earth thus the force of gravity between them is weak in comparison to the force between the earth and ourselves. The force of the gravitational pull is strong enough to keep it in orbit but not strong enough to pull it down." (grade 12)

"The moon maintains its orbit because its mass is so great it isn't pulled in. Plus the circular orbit keeps it moving in space." (grade 8)

(Although this example contains some flaws, e.g., a large mass would increase gravitational attraction and orbit does not cause the movement, the student indicates an understanding that the moon's orbit is dependent on one of the two factors—the moon's inertia.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

13% 12% **Identifies TWO OPPOSING FORCES but not correctly.**

"The moon remains away from the earth because the outer force of the universe acts upon it. The earth's gravity balances this force so that the moon remains in orbit." (grade 12)

36% 26% **GRAVITY with incorrect explanation of this one force.**

"The moon does not fall to earth because of gravity, an electromagnetic field which balances the universe." (grade 12)

5% 9% **The moon is in its orbit/place/position.**

12% 12% **Response suggests there is no gravity in space.**

12% 11% **Other incorrect response.**

5% 7% **Blank**

Comments

The lack of growth between grades eight and twelve is particularly noticeable in the responses to this question. Those students who answered gravity were on the right track; they seem to have some inkling of the relationship between gravity and motion, but could not reason the relationship between the two. Instead they arrived at a convenient explanation of why the moon didn't fall to earth due to gravitational attraction. This corresponds to research indicating that students often lack a real understanding of scientific concepts. The meaning of a term becomes attached to the particular context in which it is taught and does not transfer to other relevant situations. The contradictions that occur are often unrecognized by the student or teacher.

The principle entailed in the interaction of gravity and inertia is universal. It can be emphasized in many forms across the teaching of science, so that students who understand the principle of falling bodies on earth can apply that principle to the earth and the moon. Responses to this question imply that students do no recognize the application of the principle in this context.

ON THEIR OWN

Students' Responses to Open-Ended Questions In Math, Reading, Science, Social Studies

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. Raynolds, Jr.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Results of the 1990 Assessment

Grade Twelve

Massachusetts Department of Education
Massachusetts Educational Assessment Program

BEST COPY AVAILABLE

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Reading Background

During the past few years, the ideas that were expressed in ***Reading and Thinking*** have been refined and expanded by educators and philosophers interested in the reading process. Different themes have emerged from these activities.

One stresses the constantly shifting attitudes that the reader assumes in trying to understand any text. In contrast to the notion of reading as the accumulation of information, Judith Langer[1] proposes four different kinds of relationships that occur and reoccur during the process of reading.

Being out and stepping in. Readers use the information from the text and their background knowledge to get enough information to "step into" the author's vision. In literature, readers try to make initial acquaintance with the character, plot and setting; while in exposition they try to figure out what the topic is about.

Being in and moving through. Readers immerse themselves in the author's vision, trying to understand the meaning of the author. In exposition, readers take each new bit of information, try to understand it and to link it to what they already understand the text to have said about the topic. In fiction, readers use each new bit of information to go beyond what they already understand—asking questions about motivation, causality, and implications.

Being in and stepping out. Readers relate the text to their own knowledge and experiences. Readers of fiction use what they read in the text to reflect on their own lives, on the lives of others, or on the human condition in general. In non-fiction, readers use the text information to rethink information they already know.

Stepping out and going beyond. Readers distance themselves from the text and assume a critical stance, judging the text and relating it to other texts or experiences.

A second theme has suggested that literature is a powerful context for the teaching and learning of critical thought. As readers construct their understanding of the text, interpretations are often not possible. In fact, as Stephen Norris and Linda Phillips[2] suggest, the essence of critical reading is to raise alternative interpretations, weed out interpretations to the extent that available information will allow, and then remain with multiple possibilities. In their view, literary thinking is a complex reasoning process involving the analysis, synthesis, reformulation, linking, and generalization of ideas.

One important implication that can be drawn from this reappraisal of the reading process is that evaluation cannot be made purely on the basis of whether or not the reader's conclusions are similar to those of the teacher or test constructor. Instead, it is the quality of the reader's argument or justification that is of central importance.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

Drawing inferences proved extremely difficult for fourth graders. Given a poem, students were asked to identify the underlying theme and to draw inferences about the characters. Although the context of the poem (children walking through the woods and finding objects) appeared to be appropriate for this age, the meaning of the poem depended upon an appreciation of different perceptions that people have. In this case, the poem contrasted the imaginative, fanciful vision of the poet to the more prosaic and realistic vision of the other characters. Few of the children understood this. Among even the best readers (those at the 1550 and 1850 proficiency levels), only approximately 30 percent were able to state the underlying theme of the poem. When asked to use their interpretation to predict the action of the characters, the number dropped to 25 percent.

Students also experienced difficulty in interpreting the terms of an advertisement that was aimed at children of this age. Although 43 percent of fourth grade teachers believed that their students were well prepared to answer this type of question, this particular task was particularly challenging. Cluttered layout and deliberately ambiguous prose led the students astray. Despite the fact that the more proficient readers at this grade level showed a well-developed ability to handle details in multiple-choice questions, only approximately 30 percent were able to interpret the terms of this advertisement correctly. When asked to assume a critical stance by identifying specific examples of persuasive techniques, most students at all levels succeeded to some degree. However, only a few were able to characterize and describe the effects of such techniques on readers.

Grade Eight

The greatest discrepancy between teacher expectation and student performance occurred in response to a question which required students to identify and judge a set of assumption. Whereas less than 20 percent of teachers stated that their students were well prepared to answer this type of question, the overall success rate among students was more than twice that figure. On the other hand, performance did not follow the usual pattern of increased success rate with increased proficiency levels. Almost half of the most proficient students identified a valid assumption but, instead of judging the reasonableness of the assumption, they commented on the argument itself. Obviously, teachers did not view this question as one that fell within the area of reading instruction. They may have seen it as within the purview of social studies. Students also may have interpreted the question as one requiring argument rather than analysis. However, as discussed previously, the role of analytic and evaluative skills are not confined to a specific subject area, but are part of any type of critical reading.

Grade Twelve

On the whole, the passages on which open-ended questions were based were longer and more demanding at grade 12 than at the preceding levels; however, similar abilities were tested.

Given a passage from the James Dickey's classic *Deliverance*, students were asked to infer the character of Lewis from description and dialogue. Sixteen percent effectively identified several traits or provided an in-depth discussion of one trait, while another 21 percent gave a less comprehensive response. There was a clear distinction in the quality of responses among the different proficiency levels. Forty-two percent of students who performed at the 1850 level in the multiple-choice section of the test gave an effective and convincing response, in contrast to 24 percent of students at the 1550 level, and many fewer at the lower levels.

The ability to infer meaning from text was tested in a practical context through the use of an article on teen-age traffic deaths. The article compared a number of disparate studies on drinking and driving and concluded that more stringent laws against drunk driving may be less effective than media coverage. Students were asked to identify the main point of the article and to justify their response. This proved to be challenging for even the most proficient readers. Only 21 percent of 1850 level and 12 percent at the 1550 level could give a clear and convincing argument.

The majority of students did not make the causal connection that was suggested by the article. Many stopped at the headline or made the general statements about drinking and driving without understanding the requirements for justifying their response. Few used the statistics in the article.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.

- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.

- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of students' understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

[1] Langer, Judith. (1989) *The Process of Understanding Literature*. Center for the Learning and Teaching of Literature, 1400 Washington Avenue, University at Albany, Albany, NY 1222.

[2] Norris, Stephen and Linda Phillips. (1987) "Explanations of reading comprehension: schema theory and critical thinking theory." *Teacher College Record*. Vol. 89, 2, pp. 282-306.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Science Background

Currently, science curriculum is undergoing serious examination in light of recent findings indicating that United States students scored near the bottom in various international science exams. In response, studies (Science for All Americans, Research Within Reach: Science Education; 1994 NAEP Science Consensus) have generated curricular recommendations that are remarkably similar. Among those recommendations are:

Coverage should be in-depth.

By covering one topic thoroughly, students pick up skills that can be applied to other topics; they develop the habit of science. By stressing one concept and reinforcing that concept, students gain a depth of understanding than is far more substantial than a group of disparate facts. In other words, in the teaching of science, less is more.

Science should be taught as cross-discipline and multi-discipline.

Science is not a single area, but a series of related topics. Many of the concepts cross areas. The different subject areas do not exist in a vacuum and the interrelationship of the various disciplines provide a totality that can be offered in no other manner. Particularly at the secondary level, science is taught as series of discrete disciplines. Nor is science isolated from other disciplines. The relationship between science and mathematics is obvious, but science and the

humanities, the fine arts, and the social sciences are also intrinsically tied to one another.

Science education should begin early.

Elementary teachers might be tempted to give short shrift to science because they may feel they don't have the training to teach science, yet early science education is vital to developing students who have a love and respect for science. We know that despite much instruction, students often retain misconceptions particularly about the natural world. The earlier teachers can begin to delve into these misconceptions, the earlier they can start correcting them.

Science should begin with questions, not with answers.

Instead of teaching facts to students, a more engaging way of teaching might be to ask "Why do you think...?" A major characteristic of scientists is curiosity. This approach nurtures a sense of curiosity and the resulting sense of accomplishment that accompanies it. Answers should not be definitive but rather lead into more and deeper questions.

Science should be done, not read or heard.

Science is a series of discoveries based on careful observation and calibration. How better to actively engage students in science than with a hands-on approach, particularly in the early years. Students who begin with hands-on science develop facilities for observation, handling equipment and developing hypotheses that can be transferred to other subject areas. Doing science is more than following a pre-selected experiment with orderly steps leading to a predetermined conclusion; it is observing phenomena, developing a hypothesis, testing that hypothesis, and replicating that test in order to confirm the results.

Science should be cooperative.

Since science in the real world is a cooperative activity, it makes sense that science in school should be taught in the same manner. A scientist does not work in isolation, but shares goals and data. The scientific community tends to share knowledge and resources for the greater good of the scientific world, hence the world at large.

Science should have a practical application.

Unless students can assign a personal value to science, it remains arcane and distant. Using science helps students to appreciate its relevance in their own lives. Knowing science can only help consumers and voters make educated decisions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, students were asked to explain why or how something happened. In the area of earth science, they were asked to explain how water forms a canyon; in physical science, they were asked to explain condensation and evaporation. Among higher order thinking skills, students were asked to come up with a model of a natural phenomenon. They also had to complete the design of an experiment.

Teachers were given a selection of the open-ended questions prior to the test and asked to evaluate how well prepared their students were to answer the questions. At grade four, the relationship between actual performance on the questions and teachers' judgements of student preparedness for the questions varied. One-third of the teachers felt that their students were very well prepared to answer a question about evaporation and condensation, while eight percent of the students were able to discuss both evaporation and condensation successfully. On a question that asked students to consider characteristics of birds in relation to their habitats or food, 32% of the teachers felt that their students were very well prepared to answer. However, less than 15% of the students gave answers that were considered complete.

Grade Eight

The grade eight questions ranged from physics (laws of motion and friction and gravity) to physical science (the effects of temperature on a substance) to earth science (the causes of earthquakes) to scientific inquiry and process skills (steps in separating sand and sugar; reducing data in order to generalize;

generating an experiment). Other concepts included adaptation and the nature of sound waves.

As at grade four, teachers were given a small selection of questions and asked to evaluate their students' preparedness in responding to the task. Forty percent of the teachers felt that their students were very well prepared to describe an experiment, including the statement of hypothesis, design and necessary data. Approximately 15% of the students answered this question successfully. The largest discrepancy between teacher predictions and student performance was in response to the question "Explain why the moon does not fall to earth." Almost one-third of the teachers (31%) felt that their students were well prepared to answer this question, yet only 4% of the students could answer it successfully, with another 12% able to answer it with partial success.

Grade Twelve

At grade twelve, the questions focussed on interpreting data and making predictions including graphic representations. Other questions involved asking survey questions and generating a hypothesis. Students were also asked to explain why the moon doesn't fall to earth.

As in the case of the other grades, 12th grade science teachers were given the opportunity to consider the questions and respond to their students' preparedness. Given a scenario of fish dying unexpectedly, over one-fourth of the teachers predicted that their students were well prepared to generate two hypotheses and appropriate tests. About one-fourth of the students were able to generate good hypotheses and test them accordingly. Over one-third of teachers (35%) stated that their students were very well prepared to explain why the moon does not fall to earth, while 11% answered successfully. This question was problematic at both grades, with the teachers predicting much greater student preparedness than was the case. It may be that teachers felt they had covered the subject sufficiently.

At both grades eight and twelve, the majority of teachers (63% at grade 12 and 59% at grade 8) listed student indifference and lack of motivation as a very important reason why students do not perform well in science. And over 40% at both grades listed lack of preparation prior to entering the class as a cause of low achievement in science.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.

- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.

- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Math Background

In 1990 the National Council for Teachers of Mathematics published *The Curriculum and Evaluation Standards*, which challenged the traditional building block model of mathematics learning. According to this model, school mathematics consists of a set of discrete skills or building blocks—computational algorithms, measurement formulae, and algebraic rules. The elementary school teacher sets the first layer of blocks, the middle school teacher the second layer, and so on until, eventually the edifice is large and strong enough for students to be said to know mathematics. Although problem solving is a goal of learning, it is not actually connected to instruction. It comes later as a natural consequence of learning. The role of the teacher is to ensure that students solidify their learning at each level in order to progress to the next until, ultimately, they would be able to "do mathematics," i.e., solve real problems.

Unfortunately, for most students, this model is not working. Students are not automatically solving problems. They are not making the connections. Furthermore, most are getting stuck at the lower levels, so they are not even allowed the chance to meet real problems.

In contrast, the Standards suggested a different model for mathematics learning—one of a network of skills and concepts that are related to one another. Students can go in and out of the network at any level because the ideas are connected. Furthermore, problems form

the context of this network, because solving problems is how we begin to understand what mathematics is about.

In their publication of the *Standards*, as well as the more recent *Professional Standards for Teaching Mathematics*, the National Council for Teachers of Mathematics stresses understanding over memory, relevance over training. It advocates an expansion of the curriculum to include new topics—such as probability and statistics, patterns and relationships, discrete mathematics—that reflect the mathematical needs of society. It gives greater emphasis to other topics—such as geometry and measurement—that have often been relegated to the back of the book. It delineates four basic abilities that are essential to all topics: reasoning, communication, problem solving, and the ability to see the connections between mathematical concepts. By including evaluation, it makes the point that assessment should be viewed as an integral part of instruction.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

The open-ended questions were designed to cover different important concepts and abilities in the fourth grade mathematics curriculum. Questions concentrated on numerical operations and the ability to extract and use information presented in graphical form. There was little stress on calculation itself. Instead, students were asked to represent multiplication and division by writing stories and drawing pictures. They were required to interpret the data that were presented in a set of graphs and were asked to organize data. A final set of questions dealt with spatial visualization and geometric shapes.

When given a small selection of the actual questions, less than a quarter of the teachers replied that they believed their students to be well prepared to answer most; although another half believed that they were somewhat prepared. The exception was in response to a question asking students to write a story problem for a simple computation. In this case, 40 percent of the teachers replied that their students were well prepared.

In general, fourth graders were not hesitant to answer questions that may have been unfamiliar to them. In most cases, their answers showed a willingness to apply their reasoning and general knowledge to the task.

Grade Eight

The set of questions at this level contained both familiar word problems involving percentages and fractions and less common graphical problems that required students to extract and synthesize different types of information. As in the case of the fourth grade, teachers were presented with a selection of the problems and asked to respond in terms of their students preparation. It is not surprising that many teachers (between 45% and 71%) believed that their students were very well prepared to answer the word problems and those that probed students' understanding of numerical calculations. In contrast, only 20 percent responded that their students were prepared to use a map and a time/distance graph to describe a journey. Other questions—one concerning number patterns and one that asked students to synthesize and apply information from two graphs—were positively responded to by approximately a quarter of the teachers.

At this level, students' achievement reflected teaching more than in the earlier grade. Although few questions were answered well by more than 40 percent of the students, those tended to be the more familiar computational word problems. In response to these questions, high achievers on the multiple-choice section of the test were more successful than others. However, all students performed poorly on questions that asked them to observe, reason, and play with numbers. For example, there were few students who were able to draw inferences from graphical information or to identify more than one number pattern. In general, the eighth grade students were less prepared than their younger brothers and sisters to apply their reasoning in a mathematical context that was unfamiliar to them.

Grade Twelve

Again, at grade twelve, the preponderance of questions were in the area of numeration, with some questions focused on statistics and one which measured students' understanding of area. In most cases students were asked to explain or justify their response. Answers were judged on the quality of their response and credit was given for method in addition to the solution itself.

As was the case with the other grades, teachers were presented with a selection of questions and asked to comment in terms of their students' preparation. However, teachers who answered the twelfth grade questionnaire replied with reference to all grades at the high school level. With the sole exception of the value of 5!, no more than 20 percent felt that their students had been well prepared to answer any of this sample of questions. It seems that the written justification and reasoning required by these questions is not part of the usual high school curriculum.

Summary

From the teachers' point of view, students are not well prepared to answer questions that require them to synthesize information from various sources, to discover patterns, or to draw conclusions from data and argue their point of view. In general, the younger children—whether prepared or not—were more disposed to do so, and many performed well. As students progress through the grades, they appeared to be correspondingly less willing to attempt to answer questions that go beyond the specifics of the taught curriculum or required them to apply or discuss the concepts involved.

At grades 8 and 12 familiar questions that reflected the curriculum (e.g., the words problems at grade 8; questions that involved algebra at grade 12) were correlated with performance on the multiple-choice section of the test. The relationship between multiple-choice achievement and achievement with the less conventional questions was weaker, partially because few students at all proficiency levels performed well, partially because the explanations of the best students was not consistently better than those less capable.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.

- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.

- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- Communicate levels of student achievement throughout the state.
- Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means "to know" a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in "coming to understand" any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Social Studies Background

Currently, social studies is experiencing a "back to basics" movement stressing history, civics, and geography. This trend portends new curriculum examination and revitalization. Recent research has indicated that in elementary and middle grades, teaches make false assumptions about what their students know. Textbooks contribute to this problem by making the same assumption. In a recent article by Margaret McKeown and Isabel Beck,[1] researchers considered the background knowledge that middle school students bring to the study of history, the American Revolution in particular. They say that

> "elementary social studies textbooks seem to assume knowledge of more specific concepts than students have under control...students do seem to have a general landscape involving the country being settled, the seeking of freedom and some notions of a process...Yet this seems exactly where the texts focus their material."

This miscalculation of student knowledge calls for more complete and in-depth assessments of what significant knowledge is lacking. The longer that misconceptions continue uncorrected, the more difficult it becomes to correct them. Students will integrate knowledge to accommodate misinformation.

According to McKeown and Beck, one of the greatest challenges facing educators is finding and correcting

the confusions that young students bring to class. They suggest using a framework of the topic to organize previous knowledge and new information into a coherent whole. By presenting the relevant information in a framework, students would be better able to see the totality of the topic and evaluate whether their previous knowledge will fit in. They might be able to recognize irrelevant or inaccurate information using the framework as a sounding board.

Retention of knowledge is a problem at the secondary level. Social studies must build on the knowledge structures of the past. All this calls for a unified curriculum that considers how and when to teach materials. Walter Parker in *Renewing the Social Studies Curriculum*[2] suggests that themes in social studies can begin at the elementary grades and be reinforced across the years; that is, "content at any grade level should be presented in ways that provide, insofar as possible, a comprehensive view of a complex whole." This suggests a thematic approach without sacrificing chronology or coverage, which would satisfy those who call for social studies to have breadth over depth.

Parker also discusses attributes of a successful social studies classroom.

- Lessons have substantive coherence and continuity. Disparate facts are pulled together.
- Students are given time to think—to prepare responses to questions.
- The teacher asks challenging questions and structures challenging tasks.
- Students are able to offer explanations and reasons for their conclusions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, the open-ended questions covered process skills such as map reading and interpreting graphic representations. Another question focussed on a sense of history. Other questions involved conservation and economics.

Fourth grade teachers were given a small selection of these questions and asked to predict how prepared their students were to answer them successfully. Over half the teachers felt that their students were very well prepared to compare their own lives with those of colonial school children, while 33% of the students were able to do so. In response to how well their students were prepared to write a set of directions using a map, 31% of the teachers felt that their students were very well prepared. One-quarter of the students were able to do so.

Grade Eight

Eighth grade students were asked to interpret a political cartoon. A series of questions asked students to evaluate the quality of several arguments. Both specifics of history (students were asked to place six events on a time line) and a sense of history (students had to evaluate evidence to identify a speaker and when he lived) were covered. Students were asked to generate three courses of action and their negative impacts in response to an environmental crisis.

As in grade four, social studies teachers were given a small selection of open-ended questions and asked to evaluate how prepared their students were to answer those questions. About 15% of the teachers felt their students were very well prepared to generate three courses of action and their corresponding negative impacts; approximately 20% of the students were able to do so. When asked how well prepared their students were to explain why an argument was strong, about half the teachers felt that their students were very well prepared to do so. Eleven percent of the students were able to interpret and explain the strength of an argument.

Grade Twelve

Seniors were asked to interpret population graphs and make predictions based on those graphs. They were also asked to discuss how the policy of open immigration has affected the United States socially, politically, and economically. Another series of questions asked students to interpret two different graphs of immigration rates and discuss changes the graphs illustrated. As in grade eight, seniors were asked to interpret a political cartoon, this one about gun control.

Teachers who taught twelfth grade social studies were asked to comment on the preparedness of their students when given a sample of the questions. About one fourth (27%) of the teachers felt that their students were very well prepared to discuss the political, economic, and social effects of immigration while no more than 13% of the students were able to give complete answers in any of the categories. When given two graphs of immigration rates, one spanning 155 years, the other one year, 82% of the teachers felt their students would be able to answer why the numbers were larger in one graph than the other. Thirty-five percent of the seniors were able to answer that one graph spanned a much longer period of time than the other.

At grades eight and twelve, it appears that teachers felt their students were better prepared than they were to answer those questions that dealt with higher order thinking skills, such as evaluating arguments (grade eight) and cause and effect (effects of immigration at grade twelve).

In grades eight and twelve, the majority of teachers cited student indifference and lack of motivation as important factors in students' lack of success in social studies. Almost half the teachers at both grades cited poor preparation prior to entering the class as another important factor in students' lack of success.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.

- Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.

- Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

---

[1] McKeown, Margaret G. And Beck, Isabel L. (Winter 1990). "The Assessment & Characterization of Young Learners' Knowledge of a Topic in History." American Educational Research Journal. Vol. 27, 4, pp. 688-726.

[2] Parker, Walter C. (1991). ***Reviewing the Social Studies Curriculum.*** Alexandria, Virginia: Association for Supervision and Curriculum Development.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 12

Deliverance *is a novel by James Dickey about four suburban men who go canoeing down a wild Georgia river in search of adventure and themselves. Read the following excerpt from* **Deliverance** *and answer the question that follows it.*

Deliverance
September 14th

The pale fire on the water was not subject to the current, and this seemed wonderful to me. It played and danced where it was, an invulnerable spirt that would die. We all sat without saying anything, and I was proud of us for that, and especially proud of Lewis, who I was afraid was going to expound. I stretched out on my back, paralleling the river.

There was a darkness on my inland side when I opened my eyes; I thought I had been lying there a long time. But then something filled the space again. It was Drew with his guitar. I sat up, and the water, though it still swarmed weightlessly with the cave-images of fire, now seemed on the point of swirling them down.

Drew tuned softly, then raked out a soft chord that flowed and floated away.

"I've always wanted to do this," he said. "Only I didn't know it."

He moved up the neck, drawing out chord after chord. These built and shimmered on each other in the darkness, in lonely harmony. Then he began to pick individual notes, and put the bass under them.

"It's woods music," he said. "Don't you think so?"

"Sure do."

I loved the powerful nasal country clang, the steely humming and the strings hit like hammers on rails. Drew played deep and clean, and neither of us could have been happier. He played "Expert Town" and "Lord Bateman"; he played "He Was a Friend of Mine" and "Shaggy Dad" and Leadbelly's "Easy, Mr. Tom."

"I really ought to have a twelve-string for this one," he explained, but it sounded good anyway.

Lewis brought over the cooked steaks while Drew played, and then we ate, two little steaks apiece and

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

big wedges of cake that Lewis' wife had made. The fire was leaving us; in the river it had already died.

"You know," Lewis said, "we don't have too many more years for this kind of thing."

"I guess not," I said. "But I can tell you, I'm glad we came. I'm glad to be here. I wouldn't be anywhere else, the way I feel."

"It's true, Lewis," Bobby said. "It's all true, what you said. It's great. And I think we did real good on the river. I mean, for amateurs."

"Yeah, good enough, I reckon," Lewis said. "But I'm sure glad you and I didn't get that sluggish wood canoe turned around backward just before we hit some white water. That might have been bad."

"We didn't though," Bobby said. "And I don't think it'll happen again, do you?"

"I hope not," Lewis said.

"Well, to the sleeping bags, men," I said stretching. . . .

I stood up, finally, and creaked and stooped into the tent. I was massively tired, and hated the laces of my tennis shoes which had hardened in the water until I couldn't untie them. I pulled the shoes off by main-strength, shucked off everything else as well and got into the bag and zipped it up. Drew was still playing, out on the bank; I could hear him trying out some high minor, far away. I lay back in the soft down, crinkling into the elastic resistance of the air mattress. I snapped out the flashlight and closed my eyes.

After reading the excerpt from Deliverance, what conclusions can you draw about the type of person Lewis is? Use specific details and evidence from the passage to support your response.

Major Concepts/Abilities Tested

Attention to specific details

Recognition of ambiguity

Ability to compare and contrast

Inference about character

Although Lewis plays a relatively small role in this excerpt from Dickey's novel, there are many clues to the type of person he is — his remarks, his actions, and the reactions that he provokes in others. Unlike the other characters, who seem to be relishing their day on the rapids, Lewis's remarks are begrudging, tinged with a note of fear. He admits that he is engaged on a rare adventure, but seems unwilling to savour it as he dwells on the disaster that might have befallen. There are other clues as well. The narrator is proud that Lewis refrained from "expounding," suggesting that there is a history of complaints. On the other hand, it is also Lewis who feeds his companions with steaks and a cake that his wife has made while they enjoy their rest. Students were graded on the extent to which they recognized these traits from the clues that were given.

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 16%. **Thorough and complete responses that leave the reader with a distinct impression of the character of Lewis. These responses draw upon the entire excerpt to contrast Lewis's reactions to those of his companions. Their authors are able to cite accurate references to justify their conclusions.**

"Lewis seems pessimistic and fatalistic, dwelling more on the negative than positive. He's the one that brings up the point that they are getting older and time and nature won't allow their bodies to be so adventurous anymore. He's also the person who brings up the fact that they almost capsized their canoe and when one of his friends tries to instill confidence in their future abilities and lift spirits by saying 'we didn't though and I don't think it will happen again.' Lewis refuses to give in to optimism by responding with an 'I hope not.' "

"Lewis seems to be pessimistic, almost fatalistic about life. His first line 'we don't have too many more years for this kind of stuff' shows how he thinks a lot about death. This idea is also suggested in his next line when he points out that they would have been in trouble if they had turned the canoe around. In contrast to the other men, who enjoy life, and are very carefree, Lewis worries about everything and doesn't enjoy the beauty and adventure of the outdoors."

"Lewis is the responsible and steadfast one in the group, preparing food for Drew and the speaker, but is also one for reflection and sentiment. 'We don't have too many years for this kind of thing' he says. Lewis is a realist and his sullen nature contrasts with the carefree, adventurous natures of the other two men. Lewis feels his mortality when challenged by the rapids in the river and is not ashamed to express his fear. While the other two are slightly more invincible and not as ready to express their fear. Lewis appreciates the company he is with and the atmosphere he is in, for he realizes that the time cannot last forever and should be treasured for as long as possible."

** 11% **Good inferences and accurate references to the text, but less developed than the previous responses. The reader gets a beginning sense of the character of Lewis but the response needs expansion or greater justification from the text.**

"Lewis seems to be nervous about the trip. He was afraid that the canoe was going to trip over when they hit the white water. He makes it seem like he's an old man because he says that they don't have many more years left to do that kind of thing anymore."

"Lewis seems like he worries a lot. He also seems like he thinks he is getting old. For example he says 'You know we don't have too many years for this kind of thing.' Lewis is kind of scared too when he mentions he is glad they didn't hit some white water because it could of been dangerous. Here shows examples of why Lewis worries a lot."

"Lewis is a person who is getting old because he says they don't have too many more years for this kind of thing. He is getting old and he isn't going to be able to camp anymore. He is also afraid of dying because he say he was glad that they didn't get the canoe turned around backward. If they did that could have resulted in something bad. I think that Lewis enjoys camping but he is afraid of some of the risks and hazards that it has."

* 18% **Minimal but correct interpretation, with little or no elaborations.**

"Lewis is a home type of person with a wife. He doesn't seem like he's been out in the wild before, but he seems to enjoy it now."

"He is realistic and takes things as they come but he is not an optimist. He has a negative outlook on things. You can tell this when they were talking about the boat ride about almost getting in some danger and that they don't have that many years of doing this kind of thing left."

*** *good answer*
** *correct answer*
* *minimally correct answer*

Incorrect Answers

22% Inferences about Lewis's character that are based upon accurate but minor points in the text. By ignoring more pertinent facts to focus on less significant ones, the responses did not succeed in accurately characterizing Lewis.

"Lewis is generous and adventuresome. He shows this by giving the men steaks from home and being an amateur white water rafter. Also Lewis is an older man. We can see this when he says 'You know, we don't have too many more years for this kind of thing.' "

"One can instantly observe after reading the passage that Lewis is an optimistic person. He gives forth a positive attitude. He is enthusiastic about the men and adventure. Lewis states 'You know we don't have too many years for this kind of thing.' The narrator displays his concern for Lewis in the first paragraph, he was 'especially proud of Lewis, who I was afraid was going to expound.' This shows that Lewis is a strong minded person, a determined person. Lewis also lets his friends know that he could go on with their adventure. Lewis offers support to Drew and the narrator."

"Lewis is a person that is sweet. He brought over steaks while Drew played. He's really nice. He has a wife."

25% Inferences based on incorrect reading of the text. Sometimes it is evident that the reader is confusing Lewis with the narrator.

"Lewis seems like an adventurous person. He likes canoeing, and being out in nature. Camping, hiking, backpacking, things like that. He seems to worry about what's going to happen in the future. He seems grateful to be where he is, when he's there. He doesn't dwell on the past."

"Lewis is a calm, quiet down to earth person. He appreciates nature and being alive. He has a very positive look on life. The whole passage is full of descriptive details. I know that Lewis is an optimistic kind of person because of things like 'But I can tell you, I'm glad we came. I'm glad to be here. I wouldn't be any where else, the way I feel.' "

"Lewis seems to be a naturalist to me. He loves living in nature and living on the river. He said, 'He loved the powerful nasal country clang, the steely humming and the strings hit like hammers on nails.' 'We don't have too many more years for this kind of thing.' there it shows that he enjoys it and really doesn't want it to end. They had a fire set up with the tent and sleeping bags right on the side of the white river. His closing sentence, 'I lay back in the soft down, crinkling into the elastic resistance of the air mattress. I snapped out the flashlight and closed my eyes.' He seems to really enjoy living in the great outdoors."

5% Blank

Comments

One of the great benefits of fiction is its ability to enrich our understanding of others. Through the careful selection of dialogue, tone, and description, the novelist provides the reader with opportunities to examine the attitudes and perceptions of other human beings in the kind of detail that is impossible in everyday life. In keeping with the notion of reading as a way of making sense of the world, it is important for students to recognize the clues that authors provide in establishing the personality and motivation of their characters.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 12

Read the following preface to the book **Vietnam: There and Here**; *then answer question 3.*

Vietnam: There and Here
— Preface —

Twenty-five years ago, many Americans had difficulty locating the country on a map. Today, many Americans have difficulty locating the war in their minds.

The Vietnam War was one of the most painful events in all of American history. In 1965, when the first American combat soldiers went to war in Vietnam, Americans went to war at home. Proponents of the war — called hawks — believed that the United States, born of the desire for liberty, cannot sit by when communism denies people their liberty. Opponents of the war — called doves — protested against the war for a variety of reasons. Some believed that Vietnam was torn by a revolution or a civil war and that it is not proper for the United States to meddle in the internal affairs of any country. Others believed the war was simply unwinnable.

When American troops finally came home, in 1973, they sometimes found themselves still embattled. Veterans were often vilified by both proponents and opponents of the war, who saw them either as agents of American defeat or as agents of American immorality. In many cases the veterans were simply ignored. There were no heroes. Now, years after the war in Vietnam was ended, the war at home continues.

Many Americans have never come to a conclusion about the U.S. involvement in Vietnam. This seems strange, for never before had Americans at home been offered so much information about a war. Day after day, on television they saw footage of bombing runs, platoons moving slowly through jungles, body bags. In newspapers and magazines, they read heavily illustrated articles about corruption in the South Vietnamese government, the Tet Offensive, Vietcong assassinations. At the same time, the media were covering the war at home. Day after day, Americans saw and read about peace marches, draft dodgers, draft-card burnings. It was not long before many Americans came to regard protesters as the real enemy.

After the war, Americans still disagreed, but they did so in silence; Vietnam seemed almost a taboo subject. The renewal of discussion about the war is a sign of our health and our hope. It is clear theat the Vietnam War will haunt us until we understand it. It is also clear

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

that the United States will be a weakened giant until we develop a new sense of when and how we will wage war. Too young to have heard the debate while the war was raging in Vietnam, you now hear the debate that has revived at home. This book chronicles the war in Vietnam and the turmoil it caused at home, and outlines the larger issues involved in both. I leave to you, to reader, to decide if — and why — the war was right or wrong. You were not a part of the war in Vietnam, but you will be a part of the new debate at home.

If you were to write the five chapters of a book with the title and preface above, what would five good chapter titles be?

Major Concepts/Abilities Tested

Understanding of text

Understanding of formal features such as format

Ability to recognize logical sequence

Responses

Although calling for many of the same logical skills as outlining, this task is more difficult. The purpose of a preface is to give an impression of what will lie ahead for the reader, not to lay out the content in a logical order. As a result, students had to use the author's description to speculate on the range of topics that would probably be covered and the ways in which they could best be described. They were also required to suggest the essence of each chapter in a succinct and, possibly, intriguing manner. Finally, they had to extract the logical structure of the book that the author was describing. Each of these three dimensions—range of topics, appropriateness of language, and logical sequence—was rated separately. Although success on the dimensions were strongly related to each other, students generally found it easier to express chapter titles in a succinct and intriguing manner than to structure them in a logical sequence.

Results

	range of topics	appropriate language for title	logical sequence
A best	21%	27%	29%
B average	44%	50%	35%
C poor	31%	18%	32%

no response: 4%

Below are typical student responses that reflect different ratings on the three dimensions.

Chapter 1: Vietnam: a General Overview

Chapter 2: Controversy Over the Vietnam War

Chapter 3: Heros or Not? Veterans of Vietnam

Chapter 4: Image of Vietnam at Home

Chapter 5: A Time For Reflection: The Impact of the Vietnam War

◇ ◇ ◇

Chapter 1: The Drafting

Chapter 2: Birds at War

Chapter 3: Assassinations and Snapshots

Chapter 4: When They Came Home, (No One Wa There)

Chapter 5: The War is Over (Or Is It?)

Both sets of responses capture the main points of the preface in a succinct and engaging way although the tone differs. The language in the first set suggests a more issues-oriented approach, while the second set is more colorful. The student uses word play and provocative phrases to capture the reader's attention. Both sets cover the range of possible topics that is suggested by the author of the preface and both are structured in a logical sequence.

Both sets of responses were rated as A on all three dimensions.

Chapter 1: The Reasons Why the Americans Became the Real Enemy

Chapter 2: The Hawks vs. the Doves

Chapter 3: The Answer to Why the U.S. Meddles in Other Countries Affairs

Chapter 4: Veterans and How They Were Treated After the War

Chapter 5: The Coverage of the Vietnam War

Although this response covers some of the main points that are mentioned in the preface, it neglects others such as the aftermath of the war. Language is also somewhat weak. Not only do some of the chapter titles lack the brevity and focus that is their characteristic format, they are also ambiguous (e.g. Chapters 1 and 3). Finally, possibly because of the ambiguity, the logic of the chapter sequence appears to be weak.

This response was rated as B on all three dimensions.

Chapter 1: Vietnam Will Never End

Chapter 2: The Opinionated Conflict

Chapter 3: The Rage That Lives Within

Chapter 4: Period of Realization

Chapter 5: Understanding Their Reasons

Here, the student appears to have ignored a great deal of the content that was referred to in the preface, namely, the cause of the war, the role of the media, and the treatment of the veterans. Although the language attempts to be colorful, it is oblique. The titles give little hint of the actual content of the various chapters. Finally, there is no easily discernable logical sequence to the chapter headings.

This set of responses was rated as C on all three dimensions.

Chapter 1: The Americans Go to War in Vietnam

Chapter 2: Protesters Start a Major Upheaval in America Over the War in Vietnam

Chapter 3: Americans Soldiers Return to Hatred and Opposition

Chapter 4: The American People are Made Confused about the War by a Media Whirlwind

Chapter 5: Vietnam — the War We Must Try to Understand

The titles in this response are wordy and, in the case of Chapter 4, awkwardly constructed. On the other hand, this students covered all the topics in a logical order. **Consequently, the response was rated as an A in terms of range of topics and logical sequence and C in terms of language.**

Chapter 1: The Real Story Behind Vietnam

Chapter 2: The Proponents and Opponents

Chapter 3: How the Veterans Were Treated

Chapter 4: What Is the Conclusion?

Chapter 5: Is Anyone Going To Figure It All Out?

This response begins well but falters as it proceeds. There is some question whether the title for Chapter 1 reflects the author's tone in the preface. Despite the fact that Chapters 2 and 3 clearly cover most of the topics suggested by the author, the set of titles ignores the role of the media and gives only slight indication that the background of the war will be considered. Although the language is succinct and, in some cases colorful (e.g., proponents and opponents), it lacks clarity in other cases. For example, no indication is given of the differences between Chapters 4 and 5.

Comments

The topic interested the students; their responses indicated that they understood the author's language and had a grasp of the requirements of the task. Even those who did not perform well answered the question with some understanding. In general, students showed more involvement and performed better on this task than on a similar one that asked for an outline of instructions for gravestone rubbing.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 12

Large numbers of dead fish, all white perch, have been found along the shore of the Quabbin Reservoir. Other fish species have not been dying. Give two possible hypotheses explaining these findings. Describe how each of these hypotheses might be tested.

Major Concepts/Abilities Tested

Nature of Science

Scientific Methods

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

	Hypotheses *1st*	*2nd*	
**	8%	6%	**Good hypothesis and appropriate test with controls.**

Hypothesis

"The white perch may be reacting to something that they are eating."

How to Test

"First it should be a controlled experiment. A group of white perch must be placed in an aquarium with just fresh water and food. In another tank, water and food from the reservoir should be put in along with the fish. These fish should both be taken from the reservoir. After a few weeks time, the fish should be checked on and the data should be recorded. If the fish in the second tank die, then another test with the same water and different fish should be done."

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

(The experiment, although convoluted, adequately tests the hypothesis. It could be improved, but the student does address the issue of a controlled experiment and eventually does cover all bases.)

* 17% 12% **Good hypothesis and appropriate test with some controls.**

Hypothesis

"A corporation might be dumping something into the reservoir that only had an effect on the white perch."

How to Test

"Test the water for chemicals and if one is found, get a white perch from somewhere else along with other fish from the reservoir and test the perch along with the fish to see if that chemical kills only the perch."

(There is some control of variables here with the introduction of a white perch from somewhere else, but not enough. The water should also be controlled for by comparing it to other water.)

** *good answer*
* *minimally correct answer*

Incorrect Answers

49% 44% **Good hypothesis and logical test but lacking controls.**

Hypothesis

"The water may be polluted with a chemical that only effects that type of fish."

How to Test

"It could be tested by having some test the body of water that this fish is living in to see if there are any chemicals in the water that may affect the fish."

(This response is very similar to the previous answer but without even the introduction of a control fish.)

18% 16% **Other incorrect answer.**

9% 18% **Blank**

Comments

This question can be applied to all scientific disciplines, since it examines the concept of scientific inquiry. All scientific inquiry comprises two general aspects: the consideration of evidence based on observations and measurements; and the generation of a hypothesis based on an examination of data, combined with logic and imagination. The hypothesis suggests which data to pay attention to and which data refute it. This question tries to address the issue of selection of data. Here we are concerned with students paying attention to data that will truly measure, test, or give evidence for the hypothesis, we were less concerned about the logic of the hypothesis.

More than half the students were able to form good hypotheses which were logical, reasonable, and could be tested, but faltered when it came to testing the hypothesis. Often they did not control for all the variables or they selected the wrong variables to control for. As evidenced by the results, most often they selected a logical test with no controls. Evidence from this type of test could not conclusively confirm or invalidate a hypothesis.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 12

A local merchant in a town of 30,000 people thinks there may be a relationship between political affiliation (whether people are Democrats or Republicans) and whether they shop at her store. (Out-of-towners rarely shop there.) Describe in detail a study the merchant could conduct to test her hypothesis. Your answer should include a description of the people who would be involved in the study and the information to be collected.

Major Concepts/Abilities Tested

Scientific inquiry

Statistics

Sampling

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** 8% **Identifies the *correct four groups* (town Democrats and Republicans, shoppers Democrats and Republicans), a *sampling plan* (ask the shoppers to indicate their political affiliations) and an *analysis* (compare the results of the shoppers' responses to the affiliation of the town as a whole).**

"She could take a random sample of those who shop at her store and see what political party they are affiliated with. She should then compare her results with the percentage of people affiliated with each party in the entire town (this is at the town hall). Comparing these results will enable her to see if her hypothesis is true."

* 7% **Identifies the correct *questions* AND either a *sampling plan* OR an *analysis*.**

"At the entrance to her store, she could have someone simply standing there and

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

asking each customer their political affiliations, and then explain what the survey is for. The shop owner should then place another person at a shop on the next block which is similar to the other shop. The 'Questioner' should ask the same question and then compare the results."

(Although this response recognizes the need for a comparison to be made, it does not indicate that the political composition of the town as a whole is necessary for valid comparison.)

* 6% **Identifies the *correct groups*.**

"She could find out what political affiliation her customers have and the affiliation of all the other citizens. A simple questionnaire could be handed out to her customers to see if their shopping there had any influence on their political choice."

(This response does not address the relationship between the citizens of the town and the shoppers.)

** *good answer*
* *minimally correct answer*

Incorrect Answers

40% **Identifies the correct information but the wrong group (e.g. just ask shoppers or just the general population). It does not address all four groups.**

"She could set up a survey which is to be filled out by every customer that goes to her store. It should be comprised of questions in relationship to which political party they belong to, whether or not they are local residents. This survey should give her all the information she needs."

13% **Identifies the shoppers' affiliations indirectly.**

"If they're wearing three-piece suits and buying caviar, they're probably Republicans. If they're wearing Grateful Dead T-shirts and buying hot dogs, they're probably Democrats."

(This method only measures the prejudices of the observer.)

"The merchant could offer bumper stickers for each party. By doing this, she could see how many bumper stickers for Democrats and Republicans are taken."

(This method is more reliable in discovering political affiliation than the previous one since it requires that customers self-identify with a particular party. Still, it does not address the relationship between the shoppers and the town.)

13% **Other incorrect response.**

13% **Blank**

Comments

Although many students were able to identify the correct question e.g., "Are you a Democrat or a Republican?", most were not able to identify the correct groups, e.g., the town as a whole and the shoppers in the store. The most common response was to ask the shoppers whether they were Democrats or Republicans. This implies an inability to see the larger picture. They can make the very obvious connection between the shoppers and the store, but cannot see that, if the percentages of Democrats and Republicans shopping at the store are to have any significance, they must be compared with the percentages of Democrats and Republicans in the town.

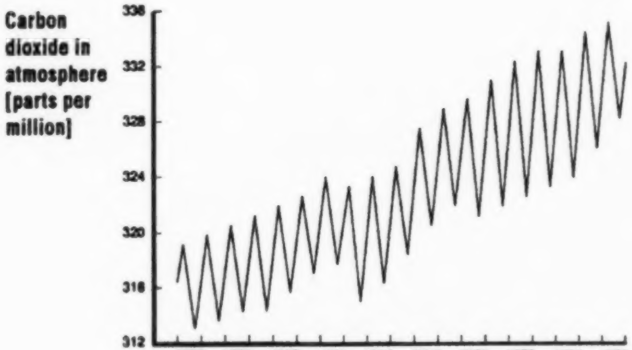
The parenthetical phrase in the question "(Out-of-towners rarely shop there.)" confused students immensely. Many suggested strategies for generating business with out-of-towners; others simply suggested ways of generating business and did not answer the question at all.

147

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 12



The graph above shows the concentration of carbon dioxide in the atmosphere over a period of almost two decades. What reason or reasons can you give for the general increase in carbon dioxide over the time period from 1958 to 1976?

Major Concepts/Abilities Tested

Environmental Issues

Impact of Technology

Reading and Interpreting Graphs

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 24% **Two or more of the following: increase in automobiles, increase in industry, increase in population, increase in use of fossil fuels, decrease in forests or any other reasonable answer with elaboration that shows some understanding of the causes.**

or

One reason with knowledgeable elaboration. Student is able to iden-

Massachusetts Educational Assessment Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

tify a cause and discuss how it is related to the effect (discusses interaction of the two).

"The population of the earth has increased considerably between 1958-1976 and human beings are great producers of carbon dioxide. The vegetation and forestation of the earth also considerably decreased 1958-1976 and plants are great consumers of carbon dioxide. The number of cars and use of the car also increased during these years, and car exhaust added to the carbon dioxide in the atmosphere. The continued burning of fossil fuels in other areas also contributed to the increase in carbon dioxide 1958-1976."

(This response gives many reasons for the increase in carbon dioxide.)

"The general increase in the concentration of carbon dioxide gas in the atmosphere over the past two decades has been a result of the development of new technology. That is, there are more cars on the road than before spewing their poisonous gases into our air. Similarly, new technology has created a greater need for industrial power — often obtained by burning coal which releases carbon dioxide. New technology has produced many new manufactured items which sometimes require, as do plastics, the combustion of certain substances in their creation."

(This response deals with one large category, "new technology," in more depth than the first answer.)

**** 32% One or more of the above reasons with some elaboration.**

"Because more people are buying and driving cars and, of course, cars blow out a lot of carbon dioxide when they burn fuel."

(Although this response gives a cause and a reason for that cause, the answer is slight.)

*** 19% Minimally correct; identifies a cause.**

"The reason for the increase in carbon dioxide from 1958 to 1976 is from the use of fossil fuels."

(This answer does not detail why the use of fossil fuels causes an increase in carbon dioxide.)

*** *good answer*
** *correct answer*
* *minimally correct answer*

Incorrect Answer

14% Incorrect explanation.

"The increase of parts per million of carbon dioxide could have been caused by the disintegrating of the ozone layer. The ozone layer exists somewhere between the surface of the earth and the sun. It essentially blocks the harmful ultraviolet rays of the sun. It is gradually disintegrating, scientists say, because of excess use of aerosol sprays such as hairspray and deodorant sprays. Thus the ozone layer can no longer absorb (as much as it used to) the carbon dioxide in the atmosphere. This effect is known as global warming."

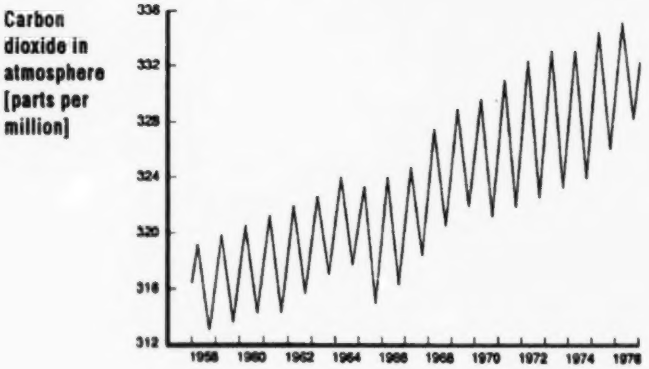
(Although the definition of the ozone layer is partially correct, the disintegration of the ozone layer did not cause the increase in carbon dioxide.)

"Plants give off carbon dioxide into the atmosphere so there has been more plants and trees in the land giving us more carbon dioxide."

10% Blank

Comments

The responses to this question reflect the fact that schools are placing a greater emphasis on environmental issues, with three-fourths of the students able to give at least a minimally correct response to this question. From this starting point, students can begin to see the relevance of science in their day to day lives. Many students and teachers consider science as strictly for scientists, when in fact, the reality of science is intimately connected with all aspects of life.



What reason or reasons can you give for the fluctuations (ups and downs) that appear in the graph?

Major Concepts/Abilities Tested

Reading and interpreting graphs

Environmental issues

Correct Answers

** 17% **Seasonal fluctuations mentioning plants and/or increase use of fossil fuels in the winter.**

"The fluctuations in the graph vary between summer and winter. In the winter, many places are heated, increasing carbon dioxide levels. In the summer, the need for heating decreases and so does the level of carbon dioxide. In the summer, the leaves on plants change carbon dioxide to oxygen. The peak of the carbon dioxide levels occurs in the winter."

* 7% **Seasonal with no explanation or partial explanation.**

"The reason on the whole why the graph is going up is obvious, because of the increase in carbon dioxide. The reason for the fluctuation is because, not only is the carbon dioxide changing during the years but during different seasons of the year, so it is constantly changing all year round."

(Unlike the first answer, this student does not explain why the seasons cause the fluctuations.)

* 4% **Measurement error or needs more information.**

"The main reason I feel there are fluctuations is because as with data from all experiments, this data is subject to experimental error. So at different times, factors could have come into play causing incorrect results. Thus the fluctuations in the graph."

(Although this answer displays an astute sense of science, the student misinterprets the effect of measurement error because with such a large quan-

tity, the effect would be slight and the effect would not have the regularity that is displayed in the graph.)

** *good answer*
* *acceptable answer*

Incorrect Answers

20% Does not recognize the regular fluctuation; instead discusses the overall trendline.

"I do not know exactly why, but if you look at the dates you can see there was a war being fought in some of the higher numbers of the graph like after 1966 when the Vietnam War was going on. We have also revolutionized new fuel burning products and we have done nuclear testing. These might be the reasons for the fluctuations in the graph."

(It's clear that the student does not address the regular fluctuations, but rather looks at the trendline and tries to attach historical events to the graph.)

"The main fluctuations of the graph might have occurred because of the wars where many people were killed. That would put less carbon dioxide into the air so the amount would go down. The reason it goes up after the down spell is because of the baby boom period after the war. As of right now, that happens to occupy the biggest age range percent of the people in the population today."

(Apparently this student took some information from a social studies graph on population by age groups, which appeared on another page of this test. Although the idea of synthesizing various pieces of information is generally useful and appropriate, in this particular situation the student selected the wrong data to use.)

25% Other incorrect answer.

"The reasons for the fluctuation is that when there is a great deal of rain and snow one year this will reduce the amount of carbon dioxide in the air."

"The reasons for the fluctuations were probably that the economy was down and the price of gas went up and therefore people's cars were being used less. That is what I would think."

27% Blank

Comments

This question focuses on two separate aspects: understanding the data presented; and applying the data to the real world.

The first aspect simply deals with the individual details in the graph. Did students read the legends and any comments about the graph? Were they able to understand the relationship in the graphs as relationships? In other words, did they understand the data simply as data? We wanted students to notice the fluctuations and consider their significance.

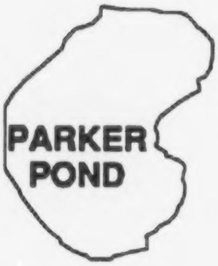
The second aspect focuses on whether they were able to correctly relate the data to the actual events represented. In other words, the students who answered that the fluctuations were seasonal paid attention to the years listed on the horizontal axis and the space between the fluctuations. By looking at both these facts, they noted the fact that the fluctuations occurred regularly within one-year periods.

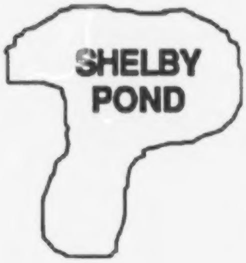
This proved to be an extremely difficult task for the majority of students. Over one quarter of the students did not even attempt the question, and more than half of those responding failed to recognize that the regularity of the fluctuations suggested a naturally occurring event.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 12





Theo wants to find out which pond covers the larger area. He does not need to know the two areas, just which is larger. Theo claims that all he has to do is measure the distance around each pond to find out what he wants. Will his method work? Make a convincing argument for your answer.

Major Concepts/Abilities Tested

Understanding of area/perimeter

Ability to communicate mathematically

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 9% **No, discusses difference between perimeter and area and gives counter-example.**

"This method will not work. What Theo is measuring is circumference, which is different from surface area. One pond may have a circumference with many bends and turns increasing the circumference, whereas the other may be more spherical with few turns, lowing the circumference. But the surface area of the pond with less circumference may be much larger than the surface area of the pond with many

Massachusetts Educational Assessment Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

bends. The two dimensions, area and circumference, are different and this should not be compared."

(This student answered the question by producing an argument why the same perimeter does not necessarily imply the same area. The student appears to recognize the requirements of a convincing argument by giving a counter-example.)

**** 8% No, states difference but no argument.**

"Theo's argument would not work. If he measured the outside of the pond by walking around it, he would get the perimeter of the ponds, not the area that they cover — which is what he needs."

(Merely defining the terms "perimeter" and "area" does not constitute an effective argument against a proposition.)

*** 17% No, minimal response without explanation.**

"Wrong. Because he wants to find which covers a larger area."

*** *good answer*
** *correct answer*
* *minimally correct answer*

Incorrect Answers

13% No, explanation unclear or incorrect.

"No, he not only needs to measure the distance around each pond but the diameter, and the water depth. After all three measurements, he would be able to make a more accurate answer."

(This student interprets Theo's argument in terms of volume, despite the fact that "covering" is a terms that refers to area, not volume.)

"He could simply measure both ponds' width and length and determine that way which is larger."

(This student is typical of many who responded with an automatic application of a formula to the word "area.")

37% Yes

"Yes, because if you measure the distance around both of the ponds and compare them, then the larger measurement will cover the larger area."

"The distance around each pond measures how large of an area it covers. So that would be the logical thing to do."

10% Blank

Comments

Vocabulary is usually cited as the reason for the notoriously poor results on area problems. Undoubtedly, a confusion in the meaning of "area" and "perimeter" accounts for some errors, but the results above point to a more profound source. When given a problem that tapped their conceptual understanding, only a third of the students were able to recognize the fact that "area" implies something more than a linear measurement. Furthermore, less than 10% were able to produce a convincing argument for their belief. The responses to this question clearly illustrate the need for discussion in mathematics. Unless students are able to examine and explore different instances of a concept, they will be reduced to a mindless application of barely remembered formulas.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 12

Roger says that raising the score on a high-scoring test paper would raise the class average on the test more than raising the score on a low-scoring paper by the same amount. Is Roger right or wrong? Use the space below to explain or prove your answer to someone who does not agree with you.

Major Concepts/Abilities Tested

Statistics

Reasoning

Communication

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** 20% **Roger is wrong with a good explanation.**

"He's wrong. As long as the number of points added was the same, it wouldn't matter. To get the class average they have to add all the scores up and then divide by the number of people in the class. The total number of all the scores will be the same no matter where the points were added."

"Raising one score or another by the same amount would increase the average equally. Since average is a computation of the total points in the test, the same average will be attained."

** 16% **Roger is wrong with a good example.**

"Either way you add the points you are going to end up with the same amount. For example, Roger's grades on two quizzes are 80 and a 20. If the teacher adds 20 points to the low grade, the grades would be an 80 and 40 for an average of 60

Massachusetts Educational Assessment Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

percent. If you add 20 points to the higher grade, you get 100 and 20 for grades. Roger's average would still be 60 percent. I feel Roger is wrong. Whichever way you add the points the final outcome will be the same."

*** 15% Roger is wrong without justification or unclear justification.**

"Wrong because the test average would be the same because they are raising the score by the same amount."

** *good answer*
* *minimally correct answer*

Incorrect Answers

There was a wide range among the incorrect responses. Generally they broke down into the following categories:

5% High scoring papers should count more.

"If you have a 50 on a test and the teacher gives everybody ten more points you would still have a failing mark. But if you have an 80 on a test and the teacher gives everyone ten more points you would have a 90 from a B to an A, but in the other case you would still have an F."

"Roger is right, because if you take a paper with a high score and you raise the score on the high scoring papers you will bring the average grade up, but if you take a low score and make it higher it really won't affect the class average much."

7% Low scoring papers should count more.

"Roger is wrong because if you raise the score on a high-scoring test paper, it will make the rest of the tests lower than if you raise the score on a lower scoring test. If the lowest mark in the class is raised, then all the other tests have to be raised also, which can only make the class average higher."

5% Cites motivational reasons.

"I think he should raise the score more on the low-scoring papers to give the students who need it more of a chance."

"I don't agree with Roger. You shouldn't raise anybody's paper. Because you shouldn't praise those who learn quicker. You should be fair about education and grades. Don't forget we are the children of the future and we all should get the grades we deserve."

19% Other

"Roger is right, it is harder to raise the higher scoring test paper than the low scoring paper."

"Depends on what kind of averages you are talking about, mode, mean or medium."

11% Blank

Comments

A variation of this question was asked in 1988. On that occasion, a quarter of the students did not attempt the question, while 48 percent were able to supply a correct response, with four percent constructing an algebraic proof. The present version elicited far fewer satisfactory responses, but more students attempted it. Its context (the raising and lowering of student scores) may have seemed more accessible and certainly diverted some students into believing that the issue was one of fairness rather than of statistics. Nevertheless, we believe that the comments written in 1988 are equally applicable to these results:

"It was obvious from their discussion that many students had only a shaky understanding of what an average means. Many more had little understanding of the hallmarks of a mathematical argument."

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 12

Major Concepts/Abilities Tested

Statistics

Percentages

In a recent survey, Americans were asked about ownership of firearms. The findings of the study were that

*** 25 percent of American families have at least one handgun;**

*** 25 percent of American families have at least one rifle; and**

*** 10 percent of American families have at least one automatic rifle.**

A reporter used the following headline on an article she wrote about the study:

MAJORITY OF AMERICAN FAMILIES OWN FIREARMS

Should the reporter's editor accept the headline as it is? Why or why not?

Correct Answers

*** 25% No, with correct reason.**

"The headline is probably false since a family who owns a handgun might also own a rifle or automatic rifle. The headline might read: Many American families own firearms."

"He should not. He has no proof that it is a majority of Americans. It is more likely that the Americans owning a firearm have more than one in their possession. Therefore the title must be renamed."

(These students correctly understood that the sets of persons in each category of gun ownership could intersect. There is no way of knowing this from the survey.)

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

Incorrect Answers

34% No, with incorrect reasoning.

"No they shouldn't print it because only 25% of families own a gun. 25% isn't a majority."

(This student does not appear to recognize that rifles can be classified as firearms.)

"No this should not be used as a headline, a survey is only a sampling of the population, therefore not everyone answered. So we do not know if the majority of American families own guns."

"No because it doesn't say how many people were asked out of how many people."

(The first student was not willing to accept the results of a survey. If reference had been made to the possibility of a sampling error, the response might have been more convincing. As it stands, it seems to imply that credible results can be obtained only by asking each individual. The second student did not seem to understand the meaning of percentages as summary statistics for "how many people were asked out of how many people.")

"I don't think that the reporter's editor should accept the headline. I feel that it is not catchy enough. The headline is exaggerated. The majority of Americans families don't own firearms."

(This response is typical of those students who objected to the word "majority.")

35% Yes

"The reporter's editor should accept the headline, because more than half of American families own firearms. I would consider 60% of American families owning guns a lot of people and I find nothing wrong with the title."

(Almost all of the students who answered that the headline should be accepted did not recognize the possibility that the same families could be represented in more than one category.)

8% Other incorrect answer.

"No it should read 'Why are the Families of America Feeling so Insecure?' and then state the statistics."

"No, it should be something to catch the readers or viewers eye. Something like, A Nation at Arms."

(These students appeared to see the question as a stylistic task.)

"Yes, he should accept this headline, because there is freedom of speech and the freedom of press."

"No, because people of the U.S. will think that it is right to own and use them at their own will."

(Students who responded in this manner, possibly misunderstanding the purpose of the task, did not address the accuracy of the headline itself.)

3% Blank/I don't know

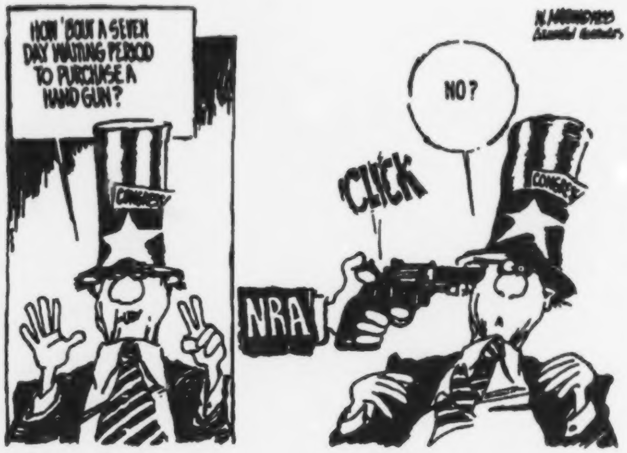
Comments

An important practical skill for any citizen is the ability to understand the statistics behind the headlines. The responses to this question suggest that only a quarter of twelfth graders are capable of doing that. Many merely added up the percentages given, without recognizing the possibility that the same families could belong to more than one category. Beyond this, however, the question elicited a wide range of responses which gave an insight into students' understanding of the role of sampling, the function of percentages, and the meaning of the word "majority." It appears that these notions, which are used to support a great deal of political and economic discussion, are not well understood by many twelfth graders.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 12


HOW 'BOUT A SEVEN DAY WAITING PERIOD TO PURCHASE A HAND GUN?
CONGRESS
NO?
CLICK
NRA
CONGRESS

What is the main idea of the cartoon?

The percentage of student responses and appropriate examples are listed under each of the categories below.

Major Concepts/Abilities Tested

Current Events

Ability to understand graphic representations

Correct Answers

(The answer should refer to three basic ideas: the issue of gun control; the power of the National Rifle Association and its position regarding gun control; the weakness of Congress in dealing with the National Rifle Association. Students should also consider the idea of humor to make a point.)

*** 6% **Answer contains all three elements. It does not necessarily have to address the issue of humor.**

"This cartoon expresses the fight for gun control. The congressman speaks about one of the options that Congress has offered. The gun held to his head represents the opposition, the National Rifle Association, who doesn't want any gun control. The congressman backs off because the

Massachusetts Educational Assessment Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

NRA is very strong and Congress is afraid of them."

**** 13% The answer contains two of the above elements.**

"Congress wants to, or is debating whether to control the use and purchase of handguns by putting a seven day delay on their purchase. However, the National Rifle Association is very much against the control of handguns, as they consider the right to bear arms a constitutional guarantee."

(This answer is very similar to the first example; however it does not address the weakness of Congress in the face of pressure from the NRA.)

*** 17% Answer contains one of the above elements.**

"The National Rifle Association doesn't want a seven-day wait period for purchasing handguns."

(This answer is marginally correct. It doesn't contain any wrong information; there is not enough of it. By identifying the NRA, the student indicates he has some sense of the cartoon's intent.)

*** *Very good answer*
** *Good answer*
* *Minimally acceptable answer*

Incorrect Answers

26% Correct information but doesn't address the question.

"The National Rifle Association fights for what's in the constitution, the right to bear arms. And they will fight for that right. We shouldn't have to change what's in the constitution."

(Although what the student says about the National Rifle Association may be true, this response does not discuss the main idea of the cartoon.)

17% Response indicates lack of prior knowledge or literal interpretation of characters.

"The man being killed by the NRA wants a seven day waiting period so he can escape."

12% Other incorrect answer

"No guns should be sold to anyone who has not a permit. Abolish guns is what they are trying to say."

9% Blank

Comments

Although students are exposed to cartoons from a very early age, many students are unable to grasp the idea of humor in the form of satire as a means of revealing some societal defect. Very few students were able to identify more than two elements of the cartoon. Because Congress and the NRA are labelled, most twelfth graders had no problems identifying the main characters of the cartoon, despite some imaginative answers (Nancy Reagan Association and National Reclamation Agency).

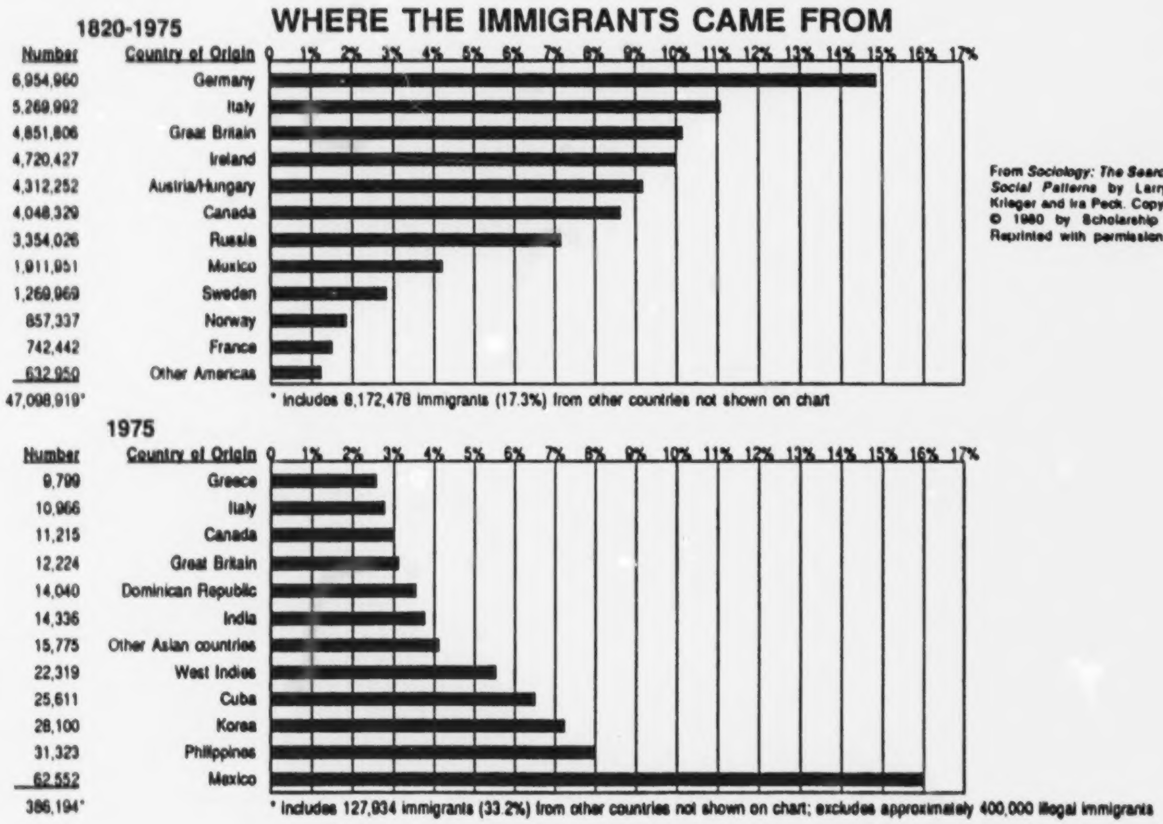
When they had to interpret the meaning, they often fell short, being unable to move from the representational to the real. Students appear to lack the ability to take the very simple situation illustrated by the cartoon and attach a larger significance to it. Just over one third of the students responded with some degree of correctness, indicating that they are not able to make the leap from the literal to the figurative. Because of their tendency toward literal interpretation, the use of irony and exaggeration only exacerbates students' difficulties.

This question also evaluated prior knowledge, which often was incorrect. Many seniors believed that the National Rifle Association was opposed to guns; others felt they sold guns.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 12



From *Sociology: The Search for Social Patterns* by Larry F. Krieger and Ira Peck. Copyright © 1980 by Scholarship Inc. Reprinted with permission.

Massachusetts Educational Assessment Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

The numbers of immigrants are given to the left of each graph. Why are there more immigrants listed on the top graph than on the bottom one?

Major Concepts/Abilities Tested

Interpretation of Graphs

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 35% Recognizes that the number of years is different.**

"The top graph spans 155 years, the bottom graph is one year."

* *correct answer*

Incorrect Answers

60% Incorrect answer

"More immigrants are listed on the top graph because the top graph is before the U.S. put low caps on immigration. Before the U.S. wanted immigrants to fill the U.S., then they became full and had to set limits."

"The illegal immigrants were excluded from the bottom chart. Also 33.2% from other countries are not shown and in the top one only 17.3% are not shown."

5% Blank

In the top graph the number of immigrants from Mexico is almost two million (1,911,951). In the bottom graph, only 62,552 are listed. Why is it that the line for Mexico is so much longer in the bottom graph than in the top one?

Correct Answers

*** 44% Answer shows the students understands that the number is a percentage.**

"The lines show percentages of total immigration. In the bottom graph, a majority (in this case 16%) of the immigrants are from Mexico."

* *correct answer*

Incorrect Answers

46% Incorrect answer

"Because in the bottom graph it says that there are approximately 400,000 aliens. I would imagine that at least one half or maybe all of them are from Mexico."

"Because it wasn't until 1975 that we let them into our country, before that we limited the amount."

10% Blank

Comments

The purpose of these two questions was to lead students to the appropriate reading of the graph. We felt that the graphs might be too difficult for students without some guidance. In the first question, students were required to recognize that the graphs dealt with different time spans; in the second, it was necessary for students to understand that the lines were percentages and not numbers.

In order to answer these two questions correctly, students had to examine the graph carefully and draw very obvious conclusions. The answers were clearly embedded in the graph. If the first skill to be used in reading a graph is correct observation of the data, then less than half the students were able to read the data correctly. Interestingly, many students mentioned the fact that 400,000 illegal immigrants were excluded, indicating that those students had read all the fine print but had not selected the appropriate facts for answering the question.

In both questions, students did not perform up to our expectations, with less than half of the students answering either question correctly. It may be that, considering the complexity of the other open-ended questions, students felt that the correct answers were too simple and, therefore, attempted to elaborate where there was no need. In response to the first question, many students attempted to explain why immigration had decreased, based on their misinterpretation of the graph. In the second question, students attempted to explain the increase of immigrants from Mexico.

The previous two questions were asked to prepare students for the next question.

What changes in the U.S. immigration pattern do the graphs show?

Major Concepts/Abilities Tested

Interpretation of graphs

Generalizing from data

Correct Answers

** 28% **Shift in immigration with generalization about trends.**

"The chart shows that between 1820 and 1975, immigration from European countries was high, but more recently, the majority of the immigrants come from Asian and Spanish countries."

* 11% **Shift in immigration with no mention of countries.**

"The countries where the least amount of people came in 1820-1975 has reversed in 1975. In 1975, the least became the most."

* 8% **Shift in immigration mentioning specific countries, not areas.**

"Fewer people are coming from Germany and Italy and more are coming from Mexico."

Incorrect Answers

4% **Increase in number.**

"That an abundance of people are starting to come to the United States more and more each year."

28% **Decrease in number.**

"Not as many people are immigrating now as there were before."

13% **Other incorrect answer**

"It shows that a lot more people are coming to this country illegally because in the bottom graph they excluded 400,000 illegal immigrants and not in the top."

"That more and more people are coming to the United States illegally and not through the immigration process. Maybe the U.S. has some tight quotas."

8% Blank

** *good answer*
* *minimally correct answer*

Comments

We asked this question to examine students' ability to generalize about immigration trends based on the information given in the graphs. We were not looking for specifics, but the recognition of trends. Students were expected to evaluate information about specific countries and to draw conclusions based on the information presented.

Just about half the students were able to recognize and describe a pattern in the immigration rates based on the evidence given in the graphs. Although calculation reveals that the number of immigrants has increased, very few students recognized this. Of those who mentioned the decrease, many tried to explain the changes as the result of tighter immigration quotas. In other words, instead of identifying the changes that occurred they attempted to answer *why* certain changes (either real or perceived) had occurred.

We can see from the responses that students do not read graphs carefully. Although students use graphs and charts at early ages, they don't seem to apply their graph reading skills in a consistent manner.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 12

Historically, the United States has had a relatively open policy toward immigration. Discuss some important ways this policy has influenced American society . . . economically, politically, socially.

Major Concepts/Abilities Tested

Approaches to history

Multiple points of view

Knowledge of immigration

Relationship of past to present

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** **Well-developed answer; each statement gives appropriate evidence and represents a balanced response.**

** 13% **Economically**

"The United States' open immigration policy led to a bigger workforce at the turn of the century. Many immigrants came to America to escape oppression and to seek employment. The industrial revolution had come and the factory system thrived on having this new source of cheap labor."

** 11% **Politically**

"The United States' open immigration policy has led to its having better relations with other countries than it might have had. Occasionally, however, the U.S. policy toward immigration has caused some tension — Chinese Exclusion Acts of the late 1800s/early 1900s for example or the

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **November 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

reluctance of the United States to accept the flood of German Jews in the 1930s and 1940s."

** 13% **Socially**

"Having an open immigration policy has led to the U.S. having a relatively more tolerant and integrated society. A great mix of nationalities populate the United States and make it a very culturally rich country. Naturally there have been problems such as the discrimination against the Irish in Boston, but on the whole, immigrants have helped the country with different foods, clothing, music and styles."

* Acceptable answer; some development to the response.

* 33% **Economically**

"Because immigration took place, many jobs were filled and little money was needed to pay workers since immigrants demanded such little pay. Companies were allowed to save and to continue to develop into industries."

(This is similar to the economic response in the first category, but does not include the historical dimension.)

* 27% **Politically**

"The influx of immigrants had allowed many different viewpoints by allowing a greater number of options and the opportunity for the American people to decide which one is right."

* 31% **Socially**

"We are expanding our culture by involving new cultures in our society. We are learning from them just as they are learning from us."

(Although they are thoughtful, the two responses above are too generalized to be considered well-developed.)

Incorrect Answers

Statement with no development.

13% **Economically**

"A lot of low paying jobs have been taken by immigrants who will work for less money when they get here."

17% **Politically**

"I believe that politically it has done some good because we believe that Americans are not made of only one race and are made of many different races and cultures, thus making us the great nation that we are."

17% **Socially**

"Immigrants may have caused some problems socially that have been brought on through no fault of their own. I speak about discrimination that 'old comers' feel for the 'newcomers.' The difference, of course, being that the old comers were here before the newcomers."

(Although there appears to be an attempt at elaboration in these responses, the elaboration only tends to restate the thesis statement.)

Answer only discusses negative aspects of influence of immigration.

22% **Economically**

"Immigrants work in low-income jobs. However, these jobs are taken from Native Americans so the rate of unemployment will increase."

(It is noteworthy that almost one-fourth of the seniors responded only negatively

when asked to discuss the economic influence of immigrants on the United States. The depressed economy seems to generate scapegoats. This student's use of the term Native American *to mean native-born people of immigrant descent is typical of many students.*)

8% Politically

"I feel that the political aspect of immigration is wrong. It's great that people can come to the United States and share our liberated freedom, but not so in the political field. Politics in our government should be run and voted upon by Americans only. Immigrants don't know enough about America to run it."

(This response reflects many that fell into this category. Students seem to feel that only those born in the United States can know about it.)

15% Socially

"I think it has done some good and some harm because we have to give immigrants a chance in improving and changing their lifestyle. It has also done some harm because now we are having racial problems. Americans are blaming the immigrants for all their problems of poverty, drugs, and even death. We have problems with Blacks against whites, Hispanics against Chinese, and everyone trying to create their own separate community because they have separate backgrounds."

(Although this student comments that immigration "has done some good" the only elaboration in the response deals with negative aspects of immigration.)

Does not address the specific area called for.

5% Economically

7% Politically

6% Socially

"Immigrants have been able to adapt easily to American culture. Jobs, housing, and communities have been kept open because of interaction between various types of people. They come to this country and we teach them our way of life."

(This student comments on how the United States has affected the immigrant rather than the effect of immigration on the United States.)

Open door immigration policy has not influenced that area of American society.

0% Economically

6% Politically

"As of today, there have been no immigrants who have run for office."

(It is only in the political category that students responded in this manner. Rather than examine one of the historical aspects of the political influence of immigrants, such as the rise of machine politics, many students simply considered that only a person born in America can run for president. There appears to be no historical consideration to this type of response.)

1% Socially

Other incorrect

7% **Economically**

"We have been able to trade freely with other nations. We are able to compensate for areas in which we are lacking (the growth of certain crops) by taking advantage of other countries abundance."

8% **Politically**

"We have not experienced a great deal of conflict with travelling through nations or passing laws between nations. Negotiations between countries have been ameliorated by open policy."

4% **Socially**

"In a way, the United Stated wants everybody to get along with everybody and sometimes it is the people in power who make everyone hate each other."

Blank

8% **Economically**

(Students had the most to say about the economic influence of immigrants, although what they say is often entirely negative.)

18% **Politically**

13% **Socially**

** *good answer*
* *acceptable answer*

General Attitude

Does the student seem to display an unbiased attitude toward immigrants?

77% **Yes**
9% **No**
13% **Can't tell**

Comments

Students were asked to generalize about the immigration experience as a whole, within the various contexts of economics, politics, and culture, and relying on their knowledge of historical and current immigrant experience.

Most students discussed the effects of immigration only on contemporary America and did not account for historical influences or aspects. Because so few students dealt with history, there is little sense that seniors view the past as having any significance to the present. For example, a large number of students focussed on African-American immigrants. While it is true that African-Americans are not native to America and that many have immigrated in the traditional sense, many do not have the same experience as European settlers to this country. Because African-Americans were not willing participants in the immigration process, their immigrant experience cannot be considered similar to the family from Germany who came to America in the late 19th century in hopes of a better life. Yet students did not focus on the historical aspect of African-American immigration; rather, they considered it to be the same as European immigration or similar to that of their own families. This inability to consider multiple perspectives often leads to simplistic understanding.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grades 8 & 12

The moon orbits the earth. Explain why the moon doesn't fall to earth.

Major Concepts/Abilities Tested

Physical science

Motion

Gravity/inertia

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

	Gr8	*Gr12*	
***	0%	3%	**Mention of vectors and resultant motions caused by inertia and gravity with explanation or drawing.**

"The moon has both angular and linear velocities. While the moon orbits the earth, the earth pulls down on it with the force of gravity. The moon fights the pull of gravity and wants to fly away from the earth. The moon stays in orbit and moves with a velocity tangent to the elliptical orbit (vector diagram included)." (grade 12)

	Gr8	*Gr12*	
**	6%	12%	**Answer discusses two dimensional motion resulting from inertia and gravity.**

"The moon is kept in orbit by the gravity of the earth and the moon wanting to go off in a straight line.

Massachusetts Educational Assessment Program

Massachusetts Department of Education **April 1991**

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas, Elizabeth Badger, and Patrick Markham.

The two opposing forces keep it in a circle around the earth." (grade 8)

(The student mentions two separate motions but does not indicate vectors. This response is weaker than one predicated on vector analysis but it does indicate a high degree of comprehension.)

* 10% 11% **Correctly describes either the effect of gravitational force or the inertial motion of the moon.**

"The moon is extremely distant from the earth thus the force of gravity between them is weak in comparison to the force between the earth and ourselves. The force of the gravitational pull is strong enough to keep it in orbit but not strong enough to pull it down." (grade 12)

"The moon maintains its orbit because its mass is so great it isn't pulled in. Plus the circular orbit keeps it moving in space." (grade 8)

(Although this example contains some flaws, e.g., a large mass would increase gravitational attraction and orbit does not cause the movement, the student indicates an understanding that the moon's orbit is dependent on one of the two factors—the moon's inertia.)

*** *very good answer*
** *good answer*
* *minimally correct answer*

Incorrect Answers

13% 12% **Identifies TWO OPPOSING FORCES but not correctly.**

"The moon remains away from the earth because the outer force of the universe acts upon it. The earth's gravity balances this force so that the moon remains in orbit." (grade 12)

36% 26% **GRAVITY with incorrect explanation of this one force.**

"The moon does not fall to earth because of gravity, an electromagnetic field which balances the universe." (grade 12)

5% 9% **The moon is in its orbit/place/position.**

12% 12% **Response suggests there is no gravity in space.**

12% 11% **Other incorrect response.**

5% 7% **Blank**

Comments

The lack of growth between grades eight and twelve is particularly noticeable in the responses to this question. Those students who answered gravity were on the right track; they seem to have some inkling of the relationship between gravity and motion, but could not reason the relationship between the two. Instead they arrived at a convenient explanation of why the moon didn't fall to earth due to gravitational attraction. This corresponds to research indicating that students often lack a real understanding of scientific concepts. The meaning of a term becomes attached to the particular context in which it is taught and does not transfer to other relevant situations. The contradictions that occur are often unrecognized by the student or teacher.

The principle entailed in the interaction of gravity and inertia is universal. It can be emphasized in many forms across the teaching of science, so that students who understand the principle of falling bodies on earth can apply that principle to the earth and the moon. Responses to this question imply that students do no recognize the application of the principle in this context.